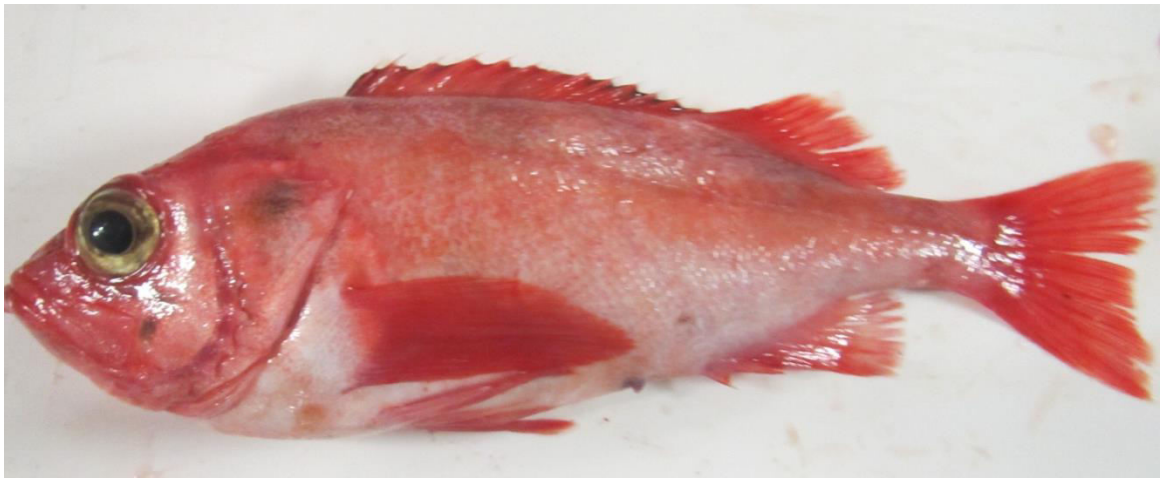
# Diversification of beaked Redfish, *Sebastes mentella*, Along a Depth Gradient in the central North Atlantic

**Peter Shum**


This thesis is presented to the School of Environment & Life Sciences, University of Salford, in fulfilment of the requirements for the degree of Ph.D.

Supervised by Prof. Stefano Mariani

March 2016



Beaked redfish *Sebastes mentella* (Travin 1951)

# Table of Contents

# List of figures

# List of tables

To my mother, Denise, without whom none of this would be possible

## Acknowledgements

I would like to express my sincere gratitude to **Stefano Mariani**, who has shown continued support throughout my research and scholarship. I appreciate his patience, encouragement and his genuine interest to catapult my career prospects. He has provided me with the freedom and opportunity to pursue my interests throughout this project while bolstering my experience in areas of bioinformatics, population genetics, marine biology and also my experience at sea. I value his comments and advice that has allowed me to mature as a researcher (and the odd game of tennis at Mersey Bowmen). It has truly been an honour to work with him and I look forward to future collaborations.

I would like to thank collaborators **Christophe Pampoulie** and **Ronald Douglas** for offering continued support, expertise and enthusiasm throughout the entire project. Chris played a vital role in organising the sampling campaign and making sure everything went smoothly. He was my first point of contact at the Marine Research Institute in Iceland and he provided any and all support when I needed it. Ron was an inspiration during this project and helped me throughout the visual studies on the retina. I learned a lot from him, especially how to "survive" at sea.

I am grateful to **Valérie Chosson**, **Carlotta Sacchi**, **Cath Hyde** and **Sarah Collins** for technical assistance during periods of field work at sea and lab work that made my life a lot easier. I would like to thank **Kiley Best** for collecting and providing tissue samples of *S. fasciatus* during offshore surveys in Canadian waters.

I owe thanks to **Andrew Mark Griffiths**, **Chrysoula Gubili**, **Marco Ciampoli**, **Sara Vandamme**, **Judith Bakker**, **Cristina Di Muri**, **Stella Minoudi**, **Ryan Joynson**, **Alice Cunnington** and **Lauren Moore** who gave me motivation during the project and helped me at various times either in the lab or crucial conversations that made my time at Salford an absolute joy.

**David Weetman** and **Adriana Adolfi** for their assistance in handling precious samples.

**Gareth Weedall**, **Steve Paterson**, **Rachel Brenchley** and the Centre for Genomic Research - University of Liverpool for technical assistance and useful advice during bioinformatics analysis. **Richard Birtles**, **Chiara Benvenuto**, **Robert Jehle**, and **Jean Boubli** for their advice and useful comments during my progression and development as a researcher.

I thank **John Hyde** and **Trevor Kenchington** for useful advice and insightful discussion on *Sebastes* during various stages of this project.

I am grateful to the crew and on board survey scientists for their gracious help during my sampling period on the Icelandic Research Vessel Árni Friðriksson – notably **Kristján Kristinsson**. The Marine Research Institute in Iceland for providing funding and facilitating transport of precious samples and the University of Salford for supporting me through the Graduate Teaching Scholarship (GTS) scheme.

Special thanks to **Valérie Chosson** from Marine Research Institute who assisted in preparing for the sampling campaign and sent me various tissues and data throughout the entire project. Her help has made this project achievable.

Finally, **Adriana Adolfi** who has been by my side from day one. She gave me the courage and motivation everyday of this Ph.D. and without her endless support I would have not made such strides as I have, for that I am grateful.

## Abstract

Despite the striking physical and environmental gradients associated with depth variation in the oceans, relatively little is known about their impact on population diversification, adaptation and speciation. The pelagic beaked redfish, *Sebastes mentella*, exhibits depth-associated patterns of population genetic substructure in the central North Atlantic, with a widely distributed population inhabiting waters between 200 and 550m depth and localised putative populations dwelling at greater depths, between 550 and 900m. This PhD project implements a multidisciplinary approach to explore and understand aspects of this biological complexity, in the context of depth.

To investigate depth as a factor driving population divergence and adaptation, sampling was carried out to target *S. mentella* at different depth layers and across most of its distribution range. Through the use of mitochondrial and nuclear markers on samples spanning a period between 2006 and 2013, the existence of two strongly divergent evolutionary lineages is revealed, with significantly different geographic distribution patterns and dwelling at different depths. The first empirical evidence to explicitly test genotype and phenotype of *Sebastes* vision associated with shifts in habitat depth is presented, by conducting spectral analysis of rod visual pigments. Additionally, *de novo* RNA-seq analysis of redfish retina transcriptome is presented for a subset of "shallow" and "deep" individuals, in order to characterise differential expression of candidate genes involved in visual perception in organisms adapted to different depths. Furthermore, a market accuracy survey of redfish sampled across Europe from various retailers was cross-referenced against data currently held in public databases. Results revealed the existence of inaccurate reference sequences in data bases, likely stemming from species misidentification from previous studies, which currently hinders the efficacy of DNA methods for the identification of *Sebastes* products. Overall, we cast new light on the role of depth in maintaining biodiversity in the oceans, and consider the practical implications of such findings.

# Chapter I

# Introduction

## 1.1 Introduction

There is increasing awareness of the vulnerability of fishery resources with over 60% of global marine populations now known to be overexploited or at their maximum sustainable limit (FAO Fisheries and Aquaculture Department, 2014). Therefore governments and intergovernmental organisations are devising management strategies aimed to improve the sustainability of fishery resources (Daw & Gray, 2005).

The scope of these measures is based on the best scientific information available on the nature of the resource. Within this framework, the discipline of fisheries biology has traditionally focused on identifying self-sustaining components by developing an understanding of the environmental and biological parameters that influence population dynamics and structure. This allows current or future population sizes to be determined and ultimately contribute to setting catch quotas (Beverton & Holt, 2012). However, failure to recognise the existence of population structure can lead to overexploitation, and consequently, have cascading impacts on the assessment and management of commercially valuable fish populations (Cadrin, Karr & Mariani, 2014).

The power to detect population structure and the degree of temporal and spatial integrity of exploited species is a difficult task. A wide variety of phenotype-based (e.g. tagging, parasite distribution, morphometrics and meristics, calcified structures) and molecular approaches (e.g. microsatellites, SNPs and mtDNA) have been used to gain valuable insight in understanding population structure, connectivity and resilience (Hughes *et al.*, 2003; Hauser & Carvalho, 2008; Cadrin *et al.*, 2010; Longmore *et al.*, 2010). In fact, the application of rapidly advancing molecular methodologies in next-generation sequencing (NGS, e.g. SNP discovery, RAD-seq & RNA-seq) have revolutionised the ability to study genome-wide signatures associated with population divergence even in the face of gene flow, and in determining the genetic basis of phenotypic traits (Davey & Blaxter, 2010; Hemmer-Hansen, Therkildsen & Pujolar, 2014), and consequently, advancing knowledge on fundamental aspects in ecology and evolution.

Research on some emblematic species, such as Atlantic cod (*Gadus morhua*) and herring (*Clupea harengus*) have played a pivotal role in the developing concepts and applications in marine stock structure analysis (McQuinn, 1997; Ruzzante *et al*., 2006; Nielsen *et al*., 2012; Hutchinson *et al*., 2015) yet, as thousands of species continue to be harvested globally, rapid and efficient population assessment are increasingly required for many other 'less traditional' fishery targets. In 1982, a pelagic fishery opened in the Irminger Sea, to target the North Atlantic beaked redfish, *S. mentella* Travin 1951, above 500m depth (commonly referred as shallow-pelagic). In the early 1990s, the pelagic fishery expanded geographically in the open Irminger Sea and moved to deeper layers, collecting fish beyond 500m (commonly referred to as deep-pelagic). This began to spark heated debates about the stock structure of *S. mentella*, which have ensued ever since (Cadrin *et al*., 2010; Makhrov *et al*., 2011; Cadrin *et al*., 2011). Initially, biologists considered that the shallow-pelagic stock could be discriminated from the deep-pelagic stock based on darker, patchy skin colour, heavier parasite infestation, and associated muscle spots, as well as a smaller size-at-maturity (Magnússon & Magnússon, 1995). However, the relationship between these stocks has been subject of strong divergence of views, especially as to whether these stocks comprise of more than one population (Saborido-Rey *et al*., 2004). Consequently an increase of studies followed with the aim of clarifying the depth-segregated population structure of *S. mentella* (Cadrin *et al*., 2010 and references therein), yet doubt remains as to the historical processes and selective forces that might have contributed to population divergence along a depth gradient in *S. mentella* in the North Atlantic (Planque *et al*., 2013).

The overarching theme of this research is to investigate the role of depth in shaping genetic and phenotypic patterns of variation in the beaked redfish, *Sebastes mentella*, in the North Atlantic. *S. mentella* is a unique system to study factors driving depth-assisted population divergence in the marine environment, yet relatively little attention has focused on mechanisms of divergence. To that end, this thesis explores the dimension of depth in facilitating population divergence.

## 1.2 Divergence along the population-species continuum

Population genetics has been intrumental in clarifying how genetic variation across a spatial continuum is distributed among different groups of individuals or "populations" and how it is influenced by environmental and demographic conditions (Hedrick, 2006).

The concept of "populations" is a core theme to the fields of ecology, evolutionary biology, and conservation biology and the criteria used to define populations fall under both evolutionary and ecological paradigms (Waples & Gaggiotti, 2006). However, fisheries science requires a higher level of ecological organisation relevent to an operational definition of a unit "stock". A stock is generally defined as groups that exhibit unique demographic dynamics to external forcing and cannot be considered exclusively as biological populations (Secor, 2013). This is because fishing and other external influences occur at different scales, such that some groups are affected more than others within a population, and the internal dynamics within a population and other groups are not necessarily homogeneous and are not exclusively internal (Secor, 2013). In practice however, it is often difficult to reach an agreement on what constitutes a stock (Gauldie, 1991) due to complex connectivity of marine populations, the wide variety of approaches, rapidly advancing methodologies, challenges in sampling and conflicting terminologies and interpretations (Cadrin, Karr & Mariani, 2014). Thus confronted with a common body of information, different researchers might come to different conclusions about the number of populations/stocks and their interrelationships.

Studies using molecular information continuously play a significant role in the assessment of exploited marine species and increasingly contribute to the development of management strategies (Hauser & Carvalho, 2008). Fisheries genetics has stimulated research to investigate the factors underpinning the dynamics and resilience of exploited species and in order to delimit the boundaries of putative fishery stocks which are the key units in the management of sustainable fisheries. DNA sequence variation is subjected to stochastic processes, such as genetic drift. For example, the DNA sequence of a specific gene (or allele) in an individual may not necessarily be the same in another individual, despite this having no effect on survival or fitness. Thus examining the homologous gene regions in different individuals can provide a measure of the frequency of allelic combinations and genetic variation both within and among individuals, populations and species. Alternatively, DNA sequence variation under the constraints of natural selection will possess some adaptive value and thereby influence an individual's fitness. These regions of the genome are underpinned by evolutionary forces (such as gene flow, genetic drift, mutation, selection and recombination) that determine phenotypic traits enabling individuals to adapt to local environments (Mariani & Bekkevold, 2013). Increasing technology-enhanced understanding of the interplay between neutral and adaptive evolutionary processes is generating a shift toward examining the genomic architecture

underpinning evolutionary important traits of different lineages (Hauser & Carvalho, 2008; Mariani & Bekkevold, 2013). One of the primary lines of interest to study adaptive variation is the ability to assess variation that underlies ecologically relevant traits (Merilä & Crnokrak, 2001), whereby neutral variation would otherwise fail to detect. This approach is particularly valuable in demographically open populations of all populations sizes, allowing the potential to detect disruptive selection, which in the long run determines adaptive divergence even in the face of gene flow (Nielsen *et al.*, 2009; Mariani & Bekkevold, 2013).

Contemporary population structure is a function of historical and current demographic processes, aided by ongoing selection pressures (Hauser & Carvalho, 2008; Nosil *et al.*, 2009a). However, different evolutionary forces can produce similar genetic patterns, making it difficult to infer which processes are responsible for the observed population structure in the absence of demographic or historical information (Strand *et al.*, 2012). Therefore it is important to gain an exhaustive picture of historical processes whereby allopatric divergence is initiated by past events, for example the Pleistocene glaciations. This period was characterized by glacial advances, colder temperatures and lower sea levels than contemporary conditions (Clark *et al.*, 2009). Consequently, populations persisting in separate glacial refugia may have diverged into genetically distinct lineages as a result of geographical isolation. Depending on the degree of reproductive isolation that evolves in allopatry, gene flow will ultimately erode divergence (Nosil *et al.*, 2009a). However, selective forces may continue to promote differentiation whereby loci under selection act on adjacent genomic regions and physically link to them (Nielsen *et al.*, 2006). Furthermore, divergent selection can promote reproductive isolation that causes barriers to overcome the homogenizing effects of gene flow (i.e. ecological speciation; Nosil *et al.*, 2009a). Hence, understanding genetic diversity of marine populations is of prime importance to help manage species persistence and its associated adaptive and evolutionary potential of populations in establising effective management strategies.

## 1.3 Environmental gradients

There has been considerable progress in identifying the scales, mechanisms and forces driving population structure in the oceans (Therkildsen *et al.*, 2013; Silva *et al.*, 2014). However, little attention focused on the mechanisms influencing deep-sea ecosystems in

the absence to obvious geographical barriers. Marine species are characterised by limited geographic barriers, often high dispersal potential, and often large effective population size (Palumbi, 1994). These attributes were originally thought limit population divergence but recent empirical work has found that gene flow is much more constrained than previously thought (Sala-Bozano *et al*., 2009; Pampoulie *et al*., 2015). A number of mechanisms have been identified that provide a platform to successfully study the interplay between gene-flow and natural selection on population structure based on geographical distance (Palumbi, 1994; Weersing & Toonen 2009), oceanographic features (Galindo *et al*., 2010; Gaither *et al.,* 2010; Selkoe & Toonen, 2011) and life-histories (Hare *et al*., 2005; Riginos *et al*, 2014). Yet, there is still a largely overlooked dimension along for which the gene flow/selection balance is likely to promote divergence in aquatic organisms. This dimension is depth.

There is accumulating evidence documenting patterns of genetic differentiation among populations separated by depth that frequently play a role in the adaptive diversification of marine fish (Karlsson & Mork, 2003; Pampoulie *et al*., 2006; Hyde *et al*., 2008). There are several reasons why variation in depth might influence population divergence. Many environmental factors - including temperature, salinity, hydrostatic pressure, scarce food resources and reduced light levels - co-vary dramatically with increasing depth (Somero, 1992; Michiels *et al*., 2008; Bergstad, 2013; Yancey *et al*., 2014), and the multidimensional nature of depth potentially pose strong divergent selection acting on depth-segregated populations over both ecological and evolutionary timescales. The multiple expanses across which individuals can disperse and populations can spread may either increase the strength of divergent selection on a particular feature of a phenotype, or it could generate 'multifarious' divergent selection acting on multiple genetically independent dimensions of a phenotype (Nosil *et al*., 2009b). Depth as a dimension for population divergence may also depend upon ontogenetic migration, spawning habitat depth, competition for resources at different trophic levels or historical events (Hyde *et al*., 2008; Vonlanthen *et al*., 2009; Stefánsson *et al*., 2009a). Thus spatial isolation by depth may result directly in reduced gene flow and the evolution of reproductive isolation between populations.

The availability of light plays a fundamental role in the behaviour and adaptation of the sensory systems of marine animals (Douglas & Partridge, 1997; Seehausen *et al*., 2008). Figure 1.1 illustrates how white light is composed of an electromagnetic spectrum of visible light or "spectral colours" of different narrow wavebands of light (Bruno &

Svoronos, 2005). In ocean water, factors such as turbidity, salinity, phytoplankton load and productivity modify ambient light spectra with depth (Thurman & Trujillo 2008). Light is reflected or refracted and is scattered or absorbed by solid particles within the water column. The visible spectrum ranges from 400-700nm at the surface and only a small fraction of visible light (1%, Michiels *et al.*, 2008) penetrates depths below 150m (Michiels *et al.*, 2008) of which most of the long wavelengths (red light; 600–700 nm) are absorbed within the first 10-15m of the water column, with ultraviolet and violet wavelengths attenuated less rapidly (Michiels *et al.*, 2008). Only short wavelengths in the blue-green spectra (400-500 nm) can be observed as deep as 1,000m (Michiels *et al.*, 2008; Warrant & Locket 2004). In the clearest ocean water, blue light (which is absorbed least) is reduced by a factor of 10 every 70m depth and all downwelling sunlight is effectively extinguished at 1,000m (Land & Nilsson, 2012). As a consequence, these processes provide a strong cline in light intensity within marine environments.

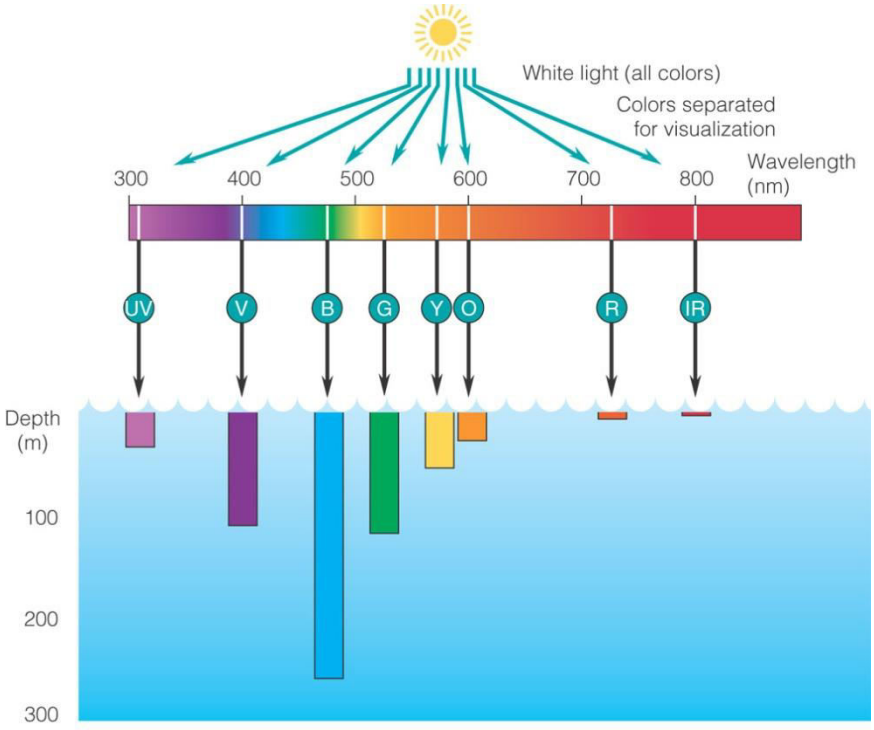



**Figure 1.1** shows the depth that light will penetrate in clear ocean water. Because red light is absorbed strongly, it has the shallowest penetration depth (10-15m), and blue light has the deepest penetration depth (1,000m, from Garrison, 2015).

## 1.4 Visual perception

There are two types of photoreceptors, rods and cones, which are distinguished by shape, the type of photopigment they contain, distribution across the retina, and the pattern of synaptic connection (Figure 1.1, Figure 1.2; Purves *et al.*, 2012). Rods and cones share the same principles of phototransduction, the cellular mechanism of light detection, and often employ homologous or sometimes even identical proteins in their phototransduction cascades. Despite their similarities, rods are highly sensitive and can detect a single photon of light (Baylor *et al.*, 1979), which makes them suited for dim light vision. On the other hand, cones are up to 100-fold less sensitive than rods (Land & Nilsson, 2012; Kefalov, 2012). Cones are capable of colour vision and have lower absolute sensitivity with high spatial acuity. There are three types of cones commonly referred to as blue or short-wavelength, green or middle-wavelength and red or long wavelength (Land & Nilsson, 2012.

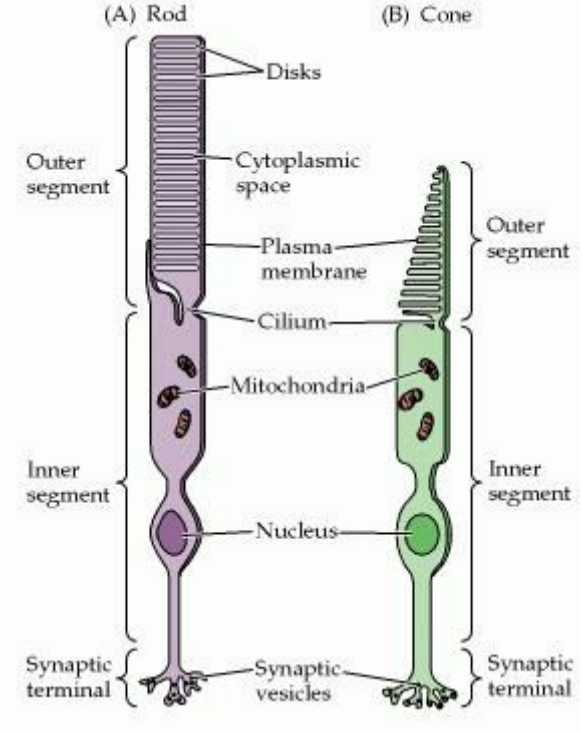



**Figure 1.2** Structural differences between rods and cones. Although generally similar in structure, rods (A) and cones (B) differ in their size and shape, as well as in the arrangement of the membranous disks in their outer segments (From Purves *et al.*, 2012).

Visual excitation is initiated in the photoreceptors of the retina when a photon is absorbed by a molecule (i.e. the chromophore) of visual pigment. The pigment is highly expressed

in specialized ciliary structures of the vertebrate rod and cone photoreceptors, called outer segments (Figure 1.2). Visual pigments are G protein-coupled receptors that consist of a protein, opsin, covalently attached to a vitamin A-derived chromophore, 11-cis-retinal (Hargrave, 2001). Upon the absorption of a photon, the chromophore is isomerized from 11-cis-retinal, to its physiologically active metarhodopsin II (meta-II) state which binds and activates transducin, the visual G-protein (Smith, 2010). In contrast to a visual excitation, the inactivation of a photoresponse to a dark-adapted state involves a two-step process: first, G protein-coupled kinases (GRKs), involved in phosphorylation, bind to active meta-II at multiple sites thereby desensitising receptor/G protein activity; in the second step, inactivation occurs when arrestin binds phosphorylated meta-II and eventually terminates the remaining activity of the opsin (see Figure 1.3; Burns & Baylor, 2001; Ebrey & Koutalos, 2001; Arshavsky *et al.*, 2002).

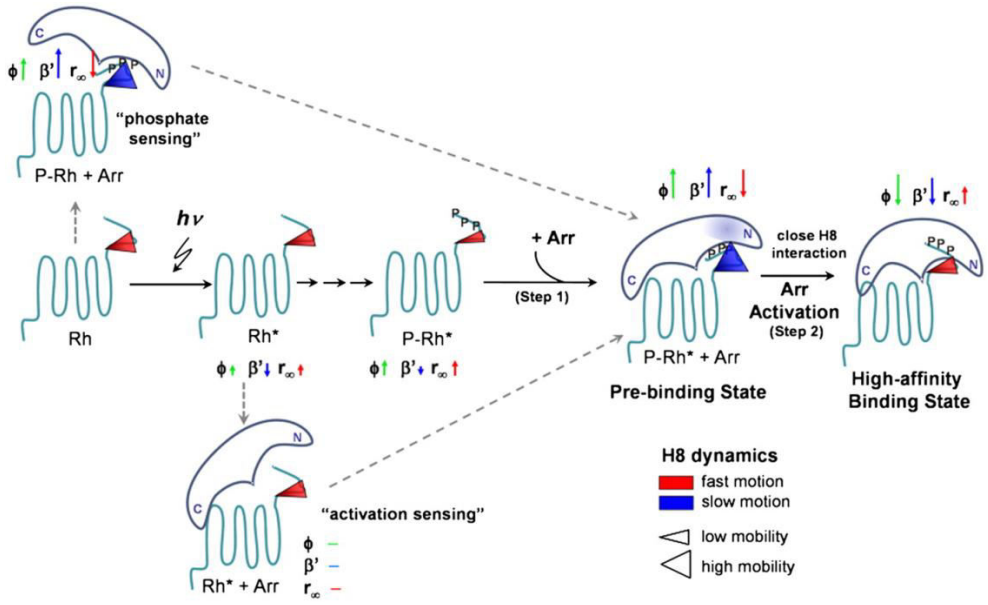



**Figure 1.3** Model for the activation of arrestin (Arr). After phosphorylation of Rh* (e.g. meta-II) by rhodopsin kinase (not shown), the prebinding and high-affinity binding state are successively formed. The conformational changes around α-helix I in the prebinding state of arrestin are indicated by a transparent blue colouring. Helical 8 (C-terminal) dynamics is visualized by the size and colour of the cone representing Helical 8 conformational space (mobility) and motion ($\phi < 2.5$ ns: red, $\phi > 2.5$ ns: blue), respectively. The dynamics changes compared to the preceding state are indicated by the corresponding anisotropy parameters φ (correlation time), β′ (conformational space/relative mobility), and r∞ (steric restriction) presented as arrows. The arrow length represents the magnitude of the change and the direction indicates the increase or decrease of the respective value (from Kirchberg *et al.*, 2011).

The light sensitivity of the visual pigments is controlled by the interaction between chromophore and opsin and is tuned to a particular wavelength of maximal absorption,

$\lambda_{max}$ (Yokoyama, 2000). In animals, visual-system rhodopsins are light-sensitive G-protein-coupled receptors (GPCRs) comprised of an aromatic cluster of amino acids located in the disc membranes of the rod outer segments and are the pigment involved in dim light vision (Smith, 2010; Yokoyama *et al*., 2008). Changes in amino acids within the transmembrane domains have been documented to alter the spectral absorbance of rhodopsin in vertebrates (Yokoyama *et al*., 2008). This raises the hypothesis that shallow-water species may show red-shift in wavelength of maximal absorbance, $\lambda_{max}$, while species that occupy progressively deeper waters are expected to show amino acid substitutions that cause blue shifts in $\lambda_{max}$ (Bowmaker *et al*, 1994). Previous studies have demonstrated cases that predict amino acid positions in the rhodopsin gene that may result in a functional role in different light environments for divergent groups of fishes (Spady *et al*., 2005; Yokoyama *et al*., 2008).

Changes in gene expression have shown to play an important role in visual sensitivity that drives adaptation in photic environments (Hofmann *et al*., 2010). Gene expression is shaped by both genetic and environmental mechanisms and thus can be considered a "molecular phenotype" (Ranz & Machado, 2006). The transcription rate of a gene can vary considerably among genotypes and/or among environmental conditions (Whitehead & Crawford, 2006; Roelofs *et al*., 2009). Consequently, gene expression might provide novel insights into population divergence because expression profiles allow the ability to uncover phenotypes, which are not immediately visible through traditional approaches such as behaviour and morphologically cryptic features (Pavey *et al*., 2010). Thus, identifying ecologically relevant expressed genes will likely enhance the efficiency of genomic approaches to address questions about ecological divergence, and ultimately speciation.

With the advent of next-generation sequencing (NGS) technologies, new sequencing-based approaches provided the ability to examine the genome-wide expression profiles in extant taxa. Whole transcriptome sequencing (RNA-seq) allows a measure of the transcript abundance of messenger RNA (mRNA) which forms the template from which proteins are translated (Wang *et al*., 2009). Thus, it provides a method to gain a detailed snapshot of transcriptome-wide gene expression of related functions from different tissues (Manousaki *et al*., 2013; Manfrin *et al*., 2015).

## 1.5 Evolutionary origins

The rockfishes of the genus *Sebastes* are the most species rich taxa among the Order Scorpaeniformes, comprising of approximately 110 species that occur in cold temperate waters throughout the North Pacific (NP, Love *et al*., 2002; Kai *et al*., 2003), North Atlantic (Nedreaas *et al*., 1994; Roques *et al*., 2002), and the Southern Hemisphere (Rocha-Olivares *et al*., 1999; Venerus *et al*., 2013). They are long-lived, viviparous fish adapted to a variety of environments, and differ in colour, ecology and behaviour (Love *et al*., 2002). The diversity of this group represents an explosive speciation event (Johns & Avise, 1998), which has been compared to the rapid radiation of cichlids in the Great East African lakes (Greenwood, 1991), resembling a species flock (Johns & Avise, 1998). Novel adaptive traits have allowed the diversifying taxa to exploit and colonise a range of ecological environments with overlapping spatial distributions (Kai *et al.* 2002; Hyde & Vetter, 2007; Sivasundar & Palumbi 2010).

The development of paleo-oceanic features conducive to the evolution and spread of *Sebastes* species in the Pacific began during the middle Miocene (~15-17 MYA) shoaling of the Indonesian Seaway (Tsuchi, 1997). This likely led to the subsequent species expansion and growth to suitable habitats and resources, while disruption of previously favourable conditions would have led to population fragmentation and extinctions (Hyde & Vetter, 2007). Developments of productive ecosystems likely prompted a process of local adaptation through niche partitioning. Thus *Sebastes* spp. colonized much of the NP as a result of a complex oceanographic temperature regime shift, which has created thermal barriers for many rockfish species (Love *et al*., 2002). As further climatic changes and expansion to new environments followed, *Sebastes* successfully colonised the North Atlantic probably as a result of the trans-arctic dispersal events approximately 3.5-4 MYA, as documented for other taxonomic groups (Dunton, 1992). It has been hypothesised that the ancestor of the *S. alutus* clade crossed the Bering land bridge into the NA as warming arctic waters allowed for the expansion and diversification into the North Atlantic (Love *et al*., 2002).

Currently, there are four species of *Sebastes* in the North Atlantic, *Sebastes viviparus* Krøyer 1845 (Norwegian redfish), *Sebastes fasciatus* Storer 1854 (Arcadian redfish), *Sebastes norvegicus* Linnaeus 1758 (golden redfish), and *Sebastes mentella* Travin 1951 (beaked redfish). These four redfish have successfully colonized a vast geographic area of the North Atlantic inhabiting depths from the subtidal down to 1,000m. *S. viviparus* is

found at much shallower depths compared to other NA redfish, between 10-120m, and has a distribution to the north-east Atlantic, from the Barents Sea to the west of Iceland, with rare occurrences off the coast of Greenland. *S. fasciatus* is restricted to the north-west Atlantic in the Gulf of Maine and along North America to the Scotian shelf and the Grand Banks from depths of 150-300m. *S. norvegicus* is most commonly found at depths less than 300-370m from the southern region of the Barents Sea, off the coast of Norway, Iceland and Greenland. *S. mentella* remains the most abundant redfish species stretching throughout the NA, from the Barents Sea, along the Norwegian coasts, around the Faroe Islands, Iceland, Greenland, the Irminger Sea and along the coasts of North America and Atlantic waters outside Canada. *S. mentella* inhabit a wide range of depths from 200m to 1,000m.

## 1.6 *Sebastes mentella*

Given the wide distribution of the pelagic beaked redfish, the spatial structure of this oceanic species has been heavily debated for decades. It was historically regarded as a single panmictic population until a pelagic fishery developed in the Irminger Sea, south-west of Iceland (Magnússon & Magnússon, 1995; Saborido-Rey *et al*., 2004). The fishery explored oceanic depth layers down to 500m before expanding geographically and to deeper depths beyond 500m. Icelandic researchers observed that fish caught at shallower layers (shallow-pelagic, <500m) could be discriminated from the deep-pelagic group based on darker patchy skin colour, heavy infestation by the parasitic copepod *Sphyrion lumpi*, differences in length-weight relationships and size-at-maturity (Magnússon & Magnússon, 1995). Researchers set out to further investigate the level of sub-structure of *S. mentella* and to determine if distinct stock units exist in the pelagic Irminger Sea, and there have been several hypothesis put forward to explain the life history of this exploited species. One hypothesis put forward agues for a "single stock hypothesis" which suggests that mature individuals of *S. mentella* comprise of a single panmictic unit that segregates by depth according to age/size across the Faroese Seas to east of Greenland. Contrary to this, some argue that there is at least two populations ("two stock hypothesis") that suggest that *S. mentella* found on the shelves (deep-demersal) and that living in deeper pelagic (deep-pelagic, >500m) waters in the Irminger Sea constitute one stock, which is separated from the "shallow-pelagic" *S. mentella* living in the upper layers (<500m) in the Irminger Sea.

Therefore, a stream of research set out to explore the depth-associated sub-structure of this oceanic species (Cadrin *et al*., 2010).

## 1.7 Genetic structure by depth

Several studies have used molecular markers for species identification and to investigate the relationships among *S. mentella* found in different areas. Genetic studies based on allozymes, mtDNA and nuclear DNA have shown significant variation between shallow- and deep-pelagic groups (Cadrin *et al*., 2010). Pampoulie and Daníelsdóttir (2008) used nine microsatellite loci on *S. mentella* and found significant allele frequency variation between shallow and deep-pelagic types as well as a significant degree of hybridization. This finding fuelled research effort to understand this complex group. Stefánsson *et al*., (2009b) carried out in-depth analysis of over 1200 genetic samples of *S. mentella* from specific locations and depths and genotyped them at 12 microsatellite loci. They revealed the presence of three distinct gene pools that show a clear segregation with depth from fish caught in the shallow zone (cluster S, <500 m), western Icelandic shelf (Cluster I, 550 m), deep sea zone (Cluster D, >500 m) (Figure 1.4). Stefánsson *et al*. (2009b) first raised the possibility of an on-going speciation event in this taxon in the Irminger Sea. They found an historical signature from microsatellite data and suggest a probable split before the last glacial maximum (~27 kyr BP), segregated in allopatry into two geographically isolated environments. Postglacial re-colonisation of the Irminger Sea (7-8 kyr BP) could have established the overlapping distribution where both groups diverged in their adaptations to local depth environments.

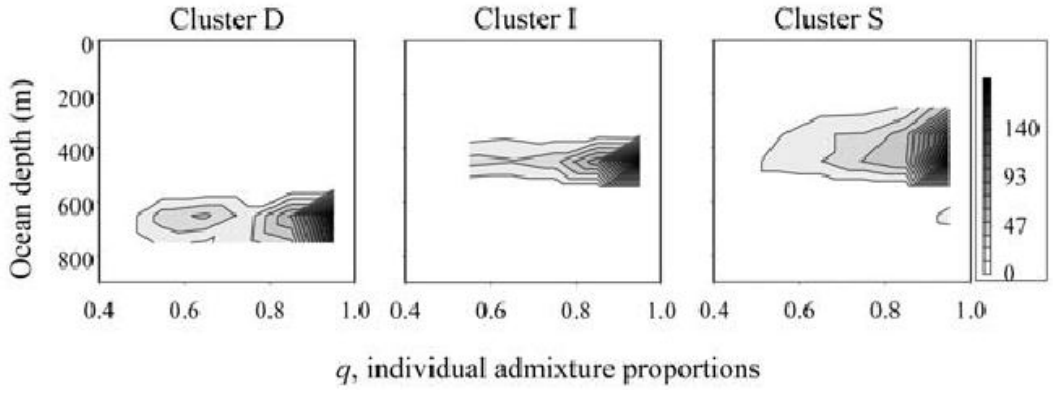



**Figure 1.4** Distribution of admixture proportions of *S. mentella* in the North Atlantic. Clusters D, I and S represent genetically distinguishable deep, demersal, and shallow *S. mentella* in the Irminger Sea and adjacent waters (from Stefánsson *et al*., 2009b).

## 1.8 Fisheries management

The population structure of North Atlantic *S. mentella* has been heavily debated for decades (Magnússon & Magnússon, 1995; Cadrin *et al* 2010). They support an important fishery industry and have become over-exploited in recent years (Marine Research Institute, 2014). Landings for the shallow-pelagic stock in the Irminger Sea and adjacent waters have been declining since 2006, effort has decreased and landings were just over 200 tonnes (t), which is the lowest they have been since fishing was first reported (Figure 1.5). In 2013, 1,500 t of the shallow pelagic redfish stock to the southeast and south of Cape Farewell, Greenland were collected. Landings for the deep-pelagic stock ranged from 75-140 kt in 1995-2004, climaxing in 1996. Since 2005, landings have decreased and were from 30-67 kt. Landings in 2013 were estimated at 45 kt, an increase of almost 13 kt over the year before and similar to those of 2011 (Figure 1.5).

Within the Irminger Sea and adjacent waters, the International Council for the Exploration of the Sea (ICES) currently recommend a separate total allowable catch (TAC) for both groups (http://goo.gl/sZfUmC; last accessed 17/12/15). ICES in 2010 advised a moratorium on the shallow stocks due to the serious decline in recent years (Figure 1.5), and the total catch limit should not exceed 20,000 tonnes to be divided among nations. Yet despite a holistic attempt to understand stock structure and identify management units (Cadrin *et al*., 2010), international agreement is often contested because some consider *S. mentella* in the Irminger Sea to be a single stock (Marine Research Institute, 2014). This consequently leads to overexploitation and has cascading impacts on the assessment and management of these commercially valuable fish stocks, which remain hotly debated (Makhrov *et al*., 2011; Cadrin *et al*., 2011). Hence, more exhaustive data are needed to fully understand the population structure in this group.

North Atlantic *Sebastes* support an important fishery industry (Figure 1.5c), yet the pattern of redfish marketed in Europe is limited and obscured by species identification made from overlapping meristic and morphological features, which leads to these species being often marketed under a single vernacular name, “redfish”. DNA based authenticity techniques have been widely used in food control and appears to be the most commonly applied approach for identifying wild fishery products (Griffiths *et al*., 2014; Mariani *et al*., 2015). For example, Logan *et al*., 2008 used a molecular approach to determine the identity of Pacific *Sebastes* (commonly known as rockfish) sold under a common name, “Pacific red snapper”. This is particularly problematic because rockfish species sold under

a common name have different conservation needs and removes the ability of consumers to choose less vulnerable species. Moreover, *Sebastes* spp. represent a recently radiating lineage and Steinke *et al.* (2009) showed that traditional molecular approaches for *Sebastes* identification is challenging because the COI fails to reliably distinguish species. Thus in order to realise the extent of *Sebastes* marketed in Europe, a product accuracy survey is required to determine the number and identity of *Sebastes* species, particularly *S. mentella* groups. Failure to recognise distinct groups within a rapidly divergent lineage will mask population declines and the risk of extinction – which potentially leads to the elimination of incipient species.

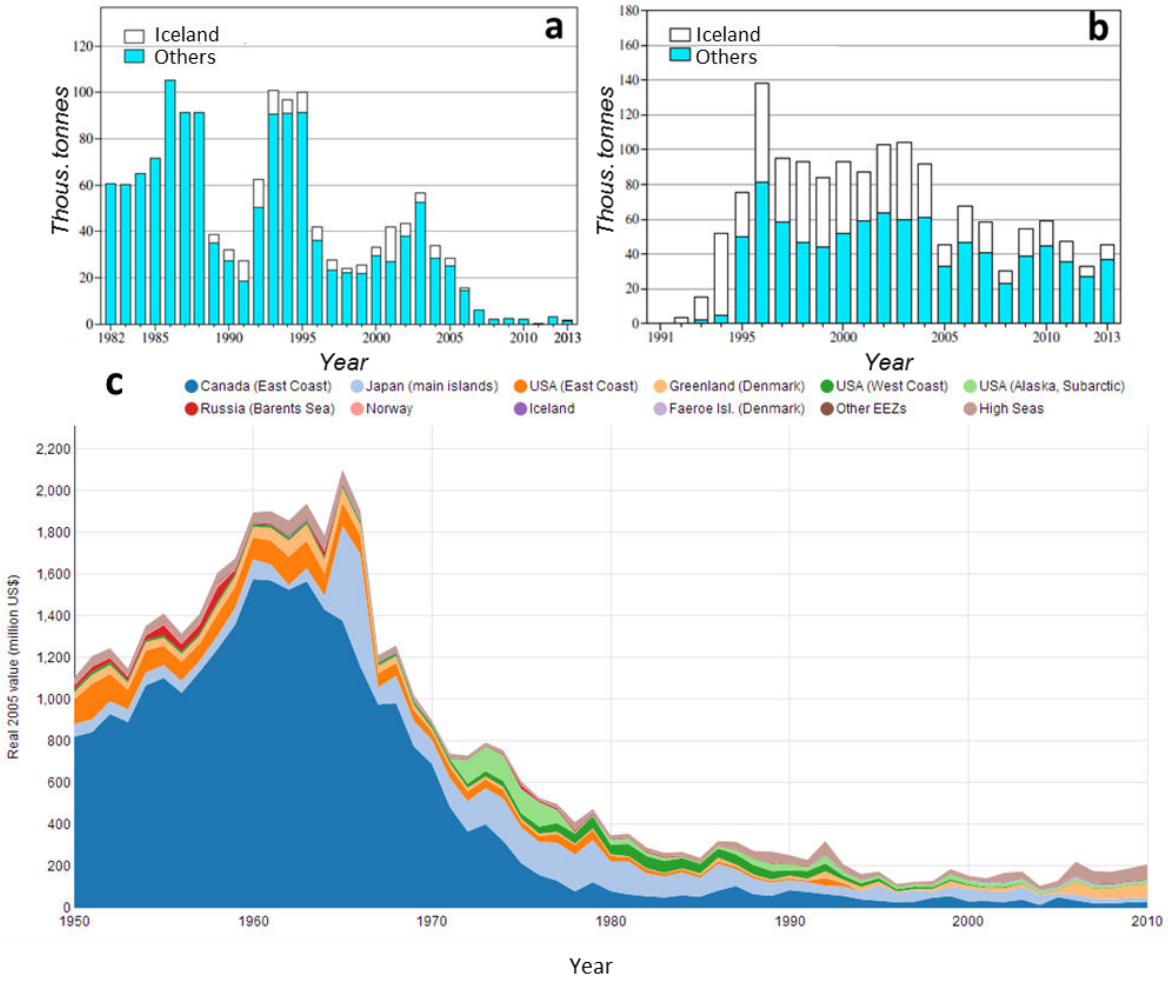



**Figure 1.5** Landings from the Irminger Sea of (a) shallow pelagic redfish since 1982 and (b) deep pelagic redfish since 1991. "Others" refer to catch data from other nations (from Marine Research Institute, 2014). (c) Real 2005 value (US$) of global catches of of Redfishes (*Sebastes*) by Exclusive Economic Zones (EEZ, From Sea Around Us http://goo.gl/boKAIn last accessed 17/02/16).

To this end a thorough examination regarding the issue of depth for this widely distributed oceanic species is needed. The ability to accurately distinguish population structure will help to understand the mechanisms underpinning their divergence along an environmental depth gradient. In order to study mechanisms of divergence, a thorough analysis of population connectivity, adaptation and genome wide variation are prerequisites for understanding population divergence, and in that context the following objectives for this thesis were identified:

- Use a multiple marker approach to resolve the interplay between longitudinal, latitudinal and vertical gradients in shaping the structure of this oceanic species in the Irminger Sea and adjacent waters (Chapter II).
- Examine the sensitivity of rod visual pigments involved in dim-light vision by conducting spectral analysis in order to link genotype with phenotype (Chapter III).
- Conduct whole transcriptome sequencing of retinal tissue from a subset of "shallow" and "deep" *S. mentella* to learn more about adaptation and plasticity of their visual systems (Chapter IV).
- Assess the state of *Sebastes* by investigating the patterns of redfish marketed from various retailers across Europe (Chapter V).

## Chapter II

# Three-Dimensional Post-Glacial Expansion and Diversification of an Exploited Oceanic Fish

## 2.1 Abstract

Vertical divergence in marine organisms is being increasingly documented, yet much remains to be done to understand the role of depth in the context of phylogeographic reconstruction and the identification of management units. An ideal study system to address this issue is the beaked redfish, *Sebastes mentella* – one of four species of "redfish" occurring in the North Atlantic – which is known for a widely-distributed "shallow-pelagic" oceanic type inhabiting waters between 250 and 550m, and a more localised "deep-pelagic" population dwelling between 550 and 800m, in the oceanic habitat of the Irminger Sea. Here, we investigate the extent of population structure in relation to both, depth and geographic spread of oceanic beaked redfish throughout most of its distribution range. By sequencing the mitochondrial control region of 261 redfish collected over a decadal interval, and combining 160 rhodopsin coding nuclear sequences and previously genotyped microsatellite data, we map the existence of two strongly divergent evolutionary lineages with significantly different distribution patterns and historical demography, and whose genetic differentiation is mostly explained by depth. Combined genetic data, analysed via independent approaches, are consistent with a late Pleistocene lineage split, where segregation by depth likely resulted from the interplay of climatic and oceanographic processes with life-history and behavioural traits. The on-going process of diversification in North Atlantic *S. mentella* may serve as an "hourglass" to understand speciation and adaptive radiation in *Sebastes* and in other marine taxa distributed across a depth gradient.

## 2.2 Introduction

While recent studies continue to amass evidence for notable patterns of population structure in the oceans (Sala-Bozano *et al.*, 2009; Therkildsen *et al.*, 2013; Silva *et al.*, 2014), the long-standing paradigm of limited geographic barriers, high dispersal potential, and large effective population size continues to hold when comparing oceanic with continental environments (Hauser & Carvalho 2008). Despite interpretive frameworks that help explain marine population structure based on geographical distance (Palumbi, 1994; Weersing & Toonen 2009), oceanographic features (Galindo *et al.*, 2010; Gaither *et al.,* 2010; Selkoe & Toonen, 2011) and life-histories (Hare *et al.*, 2005; Riginos *et al*, 2014), the reconstruction of connectivity patterns and recent evolutionary processes in oceanic species remains a difficult task, due to the afore-mentioned "marine paradigm" (see Waples, 1998; Avise, 2004; Hauser & Carvalho 2008) and the inherent difficulty – compared to terrestrial and freshwater habitats – of even locating populations and collecting representative samples over vast areas.

All of the above is further compounded by a conspicuous, yet seldom investigated, factor: depth. Depth physically multiplies the expanses across which individuals can disperse and populations can spread, over both ecological and evolutionary time scales. Depth is also associated with a wide variety of environmental factors, including temperature, salinity, hydrostatic pressure, food resources and light, which, independently or in combination, influence physiology, behaviour and adaptations (Somero, 1992; Vonlanthen *et al.*, 2008; Hyde *et al.*, 2008; Irwin, 2012; Roy *et al.*, 2012; Yancey *et al.*, 2014). As more studies document the emergence of depth-related patterns, it becomes apparent that this factor plays a role in the formation of distinct biological units, even in large, near-continuously distributed populations (Doebeli & Dieckmann, 2003; Vonlanthen *et al.*, 2008; Knutsen *et al.*, 2009; Stefánsson *et al.*, 2009a; Ingram, 2010; Roy *et al.*, 2012; Jennings *et al.*, 2013). Sea-level fluctuations during the last glacial cycle have been shown to have a role in shaping aquatic environments (Blanchon & Shaw, 1995) and are thought to have contributed to the isolation and divergence of marine organisms (Ludt & Rocha, 2015). Furthermore, the distribution of oceanic species over vast geographical areas is impacted by complex, large-scale ocean circulation processes, interplaying with bathymetric features along continental margins and mid-ocean ridges (Hemmer-Hansen *et al.*, 2007; Knutsen *et al.*, 2009). The full understanding of the role of depth in the context of physical

oceanography, complex life-histories and the historical backdrop appears as one main topical challenge in marine population biology, which may also have significant consequences for conservation and management.

One taxonomical group that has played a pivotal role in shaping our understanding of adaptive divergence along depth gradients in marine environments is the genus *Sebastes* (Alesandrini & Bernardi, 1999; Hyde & Vetter, 2007; Sivasundar & Palumbi, 2010). While most species of *Sebastes* are limited to the North Pacific (n > 100), with two occurring in the southern hemisphere, the four "young" North Atlantic (NA) species are thought to have descended from an ancestor of the North Pacific *S. alutus* lineage that invaded the NA under 3 my ago, when the Bering strait opened as a result of warming arctic waters (Raymo *et al.*, 1990; Love *et al.*, 2002; Hyde & Vetter, 2007). *Sebastes* spp. are ovoviviparous (i.e. internal fertilisation), long-lived (50-60 years), slow-growing (generation time 12-20 years) and late-maturing species with low natural mortality; such life-history characteristics influence complex population structuring (Cadrin *et al.*, 2010). The beaked redfish, *S. mentella*, widely distributed throughout the boreal waters of the North Atlantic and Arctic Oceans, exhibits complex depth-associated patterns of substructure (Stefánsson *et al.*, 2009b; Cadrin *et al.*, 2010; Shum *et al.*, 2014). While a demersal unit is recognised along the Icelandic continental shelf and slopes (Cadrin *et al.*, 2010), *S. mentella* exhibits a primarily pelagic, open-ocean behaviour, occurring in assemblages at different depth layers, with a widely distributed shallow-pelagic group inhabiting depths between 200 and 550m, and a deep-pelagic group between 550 and 800m mostly circumscribed to the Northeast Irminger Sea (Stefánsson *et al.*, 2009; Cadrin *et al.*, 2010; Planque *et al.*, 2013; Shum *et al.*, 2014).

Almost three decades of research investigating the geographic patterns and genetic structuring of *S. mentella* have resulted in intense debates (Makhrov *et al.*, 2011; Cadrin *et al.*, 2011) regarding the number of genetically distinct populations, how they are spatially structured and connected, and how and when they have a achieved their current distribution. All efforts so far had been based on traditional sampling campaigns mostly focused on geographical/fishery coverage (Cadrin *et al.*, 2010; Shum *et al.*, 2014), with a lack of adequate standardized sampling and replication across depths layers where different beaked redfish assemblages occur (Bergstad, 2013; Planque *et al.*, 2013). Furthermore, previous studies lacked either the appropriate

molecular and analytical toolkit (Stefánsson *et al*., 2009a) or the sampling coverage (Shum *et al*., 2014) to robustly test specific phylogeographic hypotheses.

Here we present the first extensive sampling, compared to previous campaigns, and phylogeographic investigation of pelagic *S. mentella* in the North Atlantic using mitochondrial and nuclear genetic markers to resolve the interplay between longitudinal, latitudinal and vertical gradients in shaping the structure of this oceanic species. Specifically, we aim to (1) examine the phylogeographic history of oceanic *S. mentella* across the North-east Atlantic, with special attention to the main two clades identified in Shum *et al*. (2014), (2) assess the degree to which mitochondrial lineage distribution fits with the currently perceived *S. mentella* stock structure, and (3) reconstruct the most realistic scenario for the evolution of this species in the North Atlantic.

## 2.3 Materials and methods

### 2.3.1 Sample collection

Samples were collected as part of the Icelandic Marine Research Institute 2013 survey, conducted from 11$^{th}$ June to the 6$^{th}$ July 2013 throughout the Irminger Sea. Trawling for redfish involved collecting replicated samples along a survey transect from three geographical zones at seven sites including shallow (above 500m) and deep (below 500m) habitats (Figure 2.1, Table 2.1). A MultiSampler unit with three separate cod-ends was attached to the end of the pelagic trawl enabling the collection of several independent discrete samples at three different depth strata in one trawl haul (Figure 2.2; Engås *et al*., 1997). The opening and closing of the nets were programmed to specific depths using a timing system with approximate sampling-depths ranging between 200-350m, 350-550m and 550-900m. These replicated sampling sites were targeted to explore the genetic variance associated with geography (latitude, longitude) and depth for *Sebastes mentella*.

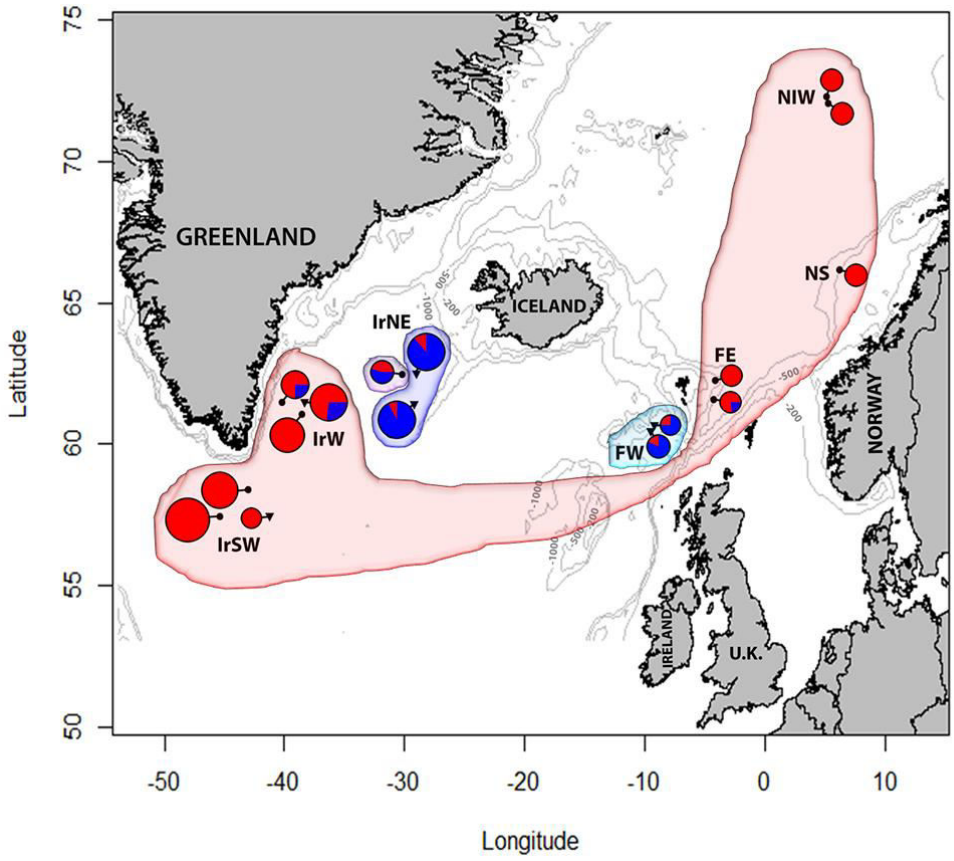



**Figure 2.1** Sample collection locations and distribution of mtDNA haploytpes. Shading show SAMOVA group structure and circles show distribution of clade A (red) and B (blue) haplogroups (see Figure 2.3). Point circles and inverted triangles indicate redfish caught above and below 500m respectively. Grey lines indicate 200, 500 and 1000 metre limit. See Table 2.1 for location codes.

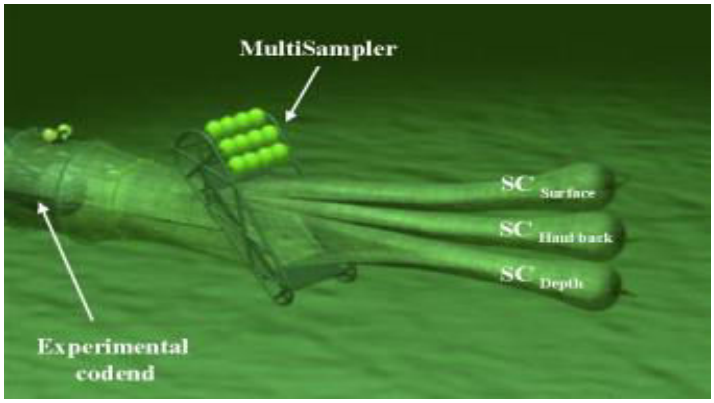



**Figure 2.2** MultiSampler unit attached to the trawl and consists of a steel frame with three codends. These can be opened and shut in order and are remotely operated through a hydro acoustic link.

**Table 2.1** Location, Group code, and year of capture, trawl depth range, position, and number of individuals (N), nucleotide ($\pi$) and haplotype ($h$) diversities of *S. mentella* investigated for 2013 and 2006/7 data from the North Atlantic.

| Location | Group code | Year | Gear | Trawl depth range | Latitude (N) | Longitude | N | $\pi$+SD | $h$+SD |
|---|---|---|---|---|---|---|---|---|---|
| **Irminger Sea** | | | | | | | | | |
| **Northeast deep** | $IrNE_{DP}13$ | 2013 | MS | 550-650 | 62.44 | -29.08 | 26 | 0.003±0.0006 | 0.754±0.090 |
| **Northeast shallow** | $IrNE_{SP}13$ | 2013 | MS | 250-300 | 62.45 | -30.21 | 11 | 0.006±0.0019 | 0.927±0.066 |
| **West shallow** | $IrW_{SP}13$ | 2013 | MS | 200-275 | 61.00 | -38.53 | 23 | 0.004±0.0007 | 0.893±0.036 |
| **West intermediate** | $IrW_{Int}13$ | 2013 | MS | 475 | 61.45 | -40.19 | 16 | 0.004±0.0004 | 0.908±0.040 |
| **West deep** | $IrW_{DP}13$ | 2013 | MS | 600 | 61.44 | -29.11 | 26 | 0.005±0.0006 | 0.898±0.029 |
| **Southwest shallow** | $IrSW_{SP}13$ | 2013 | MS | 300 | 57.44 | -45.40 | 33 | 0.004±0.0005 | 0.883±0.035 |
| **Southwest deep** | $IrSW_{DP}13$ | 2013 | MS | 550-850 | 57.45 | -41.04 | 8 | 0.004±0.0007 | 0.857±0.108 |
| **Archived samples** | | | | | | | | | |
| **Irminger sea** | | | | | | | | | |
| **Southwest shallow** | $IrSW_{SP}06$ | 2006 | Pe | 284 | 58.38 | -43 | 25 | 0.004±0.0005 | 0.883±0.035 |
| **Northeast deep** | $IrNE_{DP}06$ | 2006 | Pe | 786 | 61.44 | -29.11 | 26 | 0.002±0.0005 | 0.578±0.114 |
| **Faroe Islands West** | | | | | | | | | |
| **East** | $FW1_{DP}07$ | 2007 | Dm | 589–644 | 60.59 | -9.32 | 9 | 0.006±0.0014 | 0.917±0.092 |
| **West** | $FW2_{DP}07$ | 2007 | Dm | 704–658 | 60.44 | -9.57 | 11 | 0.006±0.0008 | 0.927±0.066 |
| **Faroe Islands east** | | | | | | | | | |
| **South** | $FE1_{SP}07$ | 2007 | Dm | 388–401 | 61.59 | -4.19 | 9 | 0.007±0.0027 | 0.833±0.127 |
| **North** | $FE2_{SP}07$ | 2007 | Dm | 366–433 | 62.21 | -4.13 | 9 | 0.005±0.0009 | 0.911±0.077 |
| **Norwegian Shelf** | $NS_{SP}07$ | 2007 | Dm | 430 | 66.14 | 6.2 | 10 | 0.002±0.0008 | 0.533±0.180 |
| **Norwegian international waters** | | | | | | | | | |
| **South** | $NIW1_{SP}07$ | 2006 | Pe | 440 | 72.01 | 5.3 | 10 | 0.003±0.0010 | 0.800±0.100 |
| **North** | $NIW2_{SP}07$ | 2006 | Pe | 445 | 72.21 | 5.15 | 10 | 0.003±0.0009 | 0.822±0.097 |

MS: multi-sampler, Dm: demersal, Pe: pelagic trawls.

Between 5-30 specimens were collected and sampled from each haul section, for a total of 143 dorsal fin clip tissue samples of *S. mentella* collected at 7 sites distributed throughout the Irminger Sea and stored in 100% ethanol. The data set was also complemented by the 50 mtDNA control region and 22 rhodopsin sequences that first revealed the existence of two depth-associated *S. mentella* clades in the Irminger Sea (Shum *et al.,* 2014; GenBank accession numbers mtDNA: KM013849-KM013898; rhodopsin: KM013899.1-KM013920.1) and by 68 randomly selected archived tissue samples (previously genotyped at 12 microsatellite loci by Stefánsson *et al*., 2009b), to increase cover the current biogeographical spread of the species: west Faroe (n=20) and east Faroe (n=18) collected off the Faroese shelf in 2007, Norwegian Shelf (n=10), and pelagic collections from Norwegian International Waters (n=20), also from 2006-2007 surveys (Figure 2.1, Table 2.1). Additional archived tissue samples of 8 *S. norvegicus* and 7 *S. viviparus* were also analysed for downstream analysis as outgroups, as well as 36 tissue samples of *S. fasciatus* collected on board the Irish research vessel *RV Celtic Explorer* in May 2012 off Newfoundland, Canada. The sampling scheme and sample sizes per site are summarized in Table 2.1.

### 2.3.2 Generation of molecular data

DNA was extracted using a modified salt extraction protocol (Miller *et al*., 1988). An approximately 500 base pair (bp) fragment of the non-coding mitochondrial control region was amplified using primers developed by Hyde & Vetter (2007; D-RF: 5′-CCT GAA AAT AGG AAC CAA ATG CCA G-3′ and Thr-RF: 5′-GAG GAY AAA GCA CTT GAA TGA GC-3′.) and sequenced using Thr-RF. The intron-free rhodopsin gene was used to amplify an approximately 800 bp fragment in 138 *S. mentella* samples using primers developed by Chen, Bonillo & Lecointre (2003); Rh193: 5′-CNT ATG AAT AYC CTC AGT ACT ACC-3′ and Rh1039r: 5′-TGC TTG TTC ATG CAG ATG TAG A-3′ and sequenced using Rh193. PCR reactions and temperature profiles followed Shum *et al.* (2014). PCR products were purified and sequenced by Source BioScience LifeSciences (http://www.lifesciences.sourcebioscience.com/). Resulting sequences were manually checked for quality, edited and trimmed using Chromas Lite 2.1.1 (http://technelysium.com.au/?page_id=13) and aligned using MUSCLE v3.7 (Edgar, 2004).

### 2.3.3 Data analysis

Mitochondrial DNA variation and structure

Molecular diversity indices, including nucleotide ($\pi$) (Nei, 1987) and haplotype ($h$) (Nei & Tajima, 1981) diversities, were estimated using DnaSP v5.10 (Librado & Rozas, 2009). Haplotype genealogies were constructed in the program HapView, following a method described by Salzburger *et al.* (2011), based on a maximum likelihood tree implemented in PHYLIP v3.695 (Felsenstein, 1989) for mtDNA. To gauge the level of population differentiation among collections, ARLEQUIN v.3.5.1.2 (Excoffier & Lischer, 2010) was used to calculate pairwise population $\Phi_{ST}$ and $F_{ST}$, with significance of pairwise differences at the level of 0.05 assessed with 10,000 permutations. As a correction for multiple tests, *p*-values were adjusted according to the modified false discovery rate method (Narum, 2006). The relationship among all sample collections were visualised by multi-dimensional scaling (MDS) of pairwise $\Phi_{ST}$ and $F_{ST}$ using the MASS (Venables & Ripley, 2002) package in the R programming environment (Team RDC, 2005). Spatial analysis of molecular variance (SAMOVA 1.0) (Dupanloup *et al.*, 2002) was used to identify distinct groups of populations within which ($F_{SC}$) genetic variance is minimal and among which ($F_{CT}$) it is greatest. This method partitions populations into a specified number of groups and the partition scheme (k) that maximizes differences between $F_{SC}$ and $F_{CT}$ is selected. The optimum number of groups was determined by running SAMOVA with two to eight groups with 100 annealing steps for each run.

*Distribution of geographically restricted alleles SAShA* – We tested the extent to which haplotypes are randomly distributed in the NA, as implemented in the program SAShA (Kelly *et al.*, 2010). Assuming haplotypes are identical-by-descent, non-random distribution of haplotypes can indicate departures from panmixia and occurrences of the same haplotypes in different locations can be considered evidence of gene flow. SAShA generates observed distribution (OM) of geographic distances among instances of each haplotype which is compared to the null distribution (EM; i.e. allele distribution under panmixia) generated from the same data. An OM significantly less than EM indicates that haplotypes are under distributed, and that gene flow is somewhat restricted. For mtDNA, we tested for significant deviation of the arithmetic mean of OM from EM (D$g$) using 1000 nonparametric permutations of the haplotype-by-location and -depth data set.

#### *2.3.3.1 Demographic changes*

First, Tajima's $D_T$ (Tajima, 1983) and Fu's *Fs* (Fu, 1997), calculated in DnaSP, were tested to examine deviations of the mitochondrial site frequency spectrum expected under the neutral expansion model for major clades (see results). Significant and negative $D_T$ and *F*s values indicate population size changes or directional selection (i.e., selective sweep; Aris-Brosou & Excoffier, 1996). Mismatch distribution analysis was also carried out, as implemented in ARLEQUIN, whereby a unimodal frequency distribution of pairwise haplotype differences is expected for populations that have experienced recent demographic expansion, and a multimodal frequency distribution is expected for populations at equilibrium. To estimate the divergence time, we used the formula $T = D_a/2\mu$, where $2\mu$ represents a general mtDNA evolutionary rate, commonly assumed to be around 11% per million years for fish mtDNA control region (Patarnello *et al.*, 2007).

To further assess the demographic changes in effective population size since the time of most recent common ancestor (TMRCA) of the identified lineages, Bayesian skyline plots (BSPs) were generated using the software package BEAST v. 1.8 (Drummond *et al.*, 2012). Analysis was performed using the best fit model of nucleotide substitution (GTR+$\Gamma$) implemented using Modeltest3.7 (Hasegawa *et al.*, 1985; Posada & Crandall, 1998) and a fixed clock was set using 11%/million years (Myr) (Patarnello *et al.*, 2007). For each analysis the Monte Carlo Markov Chain (MCMC) was set at 150 million steps, which yielded effective sample sizes (ESS) of at least 200. Once the appropriate mixing and convergence was met, the first 10% of the posterior was discarded and the remainder combined for parameter inferences. BSPs were estimated in Tracer v. 1.6 (Rambaut *et al.*, 2014) and plotted using the upper 95% highest posterior density.

#### *2.3.3.2 Marker comparison and Approximate Bayesian Computation*

To gain an exhaustive picture of genetic variation, we also investigated the nuclear intron-free, gene coding for the rhodopsin pigment, which Shum *et al.* (2014) reported to also exhibit alternative genotypes associated with depth. Haplotype genealogies for 160 *S. mentella* rhodopsin sequences were constructed in HapView using a maximum likelihood tree generated using PHYLIP. MtDNA haplogroup distributions and rhodopsin genotypic frequencies for the shallow- and deep-caught *S. mentella* groups were compared using Chi-

square tests (2x2 contingency tables). Furthermore, we re-examined microsatellite data at 12 loci, previously genotyped by Stefánsson *et al.* (2009b), for the 118 historical samples from 2006-2007 (Stefánsson *et al.*, 2009b), and reanalysed them in combination with the newly-generated mtDNA sequences. Population structure was assessed by calculating pairwise genetic differentiation through Slatkin's $R_{ST}$ and Weir & Cockerham's $F_{ST}$, with 9,999 permutations carried out to obtain significance levels, using GenAIEx 6.501 (Peakall & Smouse, 2006). $F_{ST}$ measures genetic differentiation based on allele identity whereas its analogue $R_{ST}$ is an allele size measure of differentiation that assumes a strictly stepwise mutation process. $R_{ST}$ is expected to be larger than $F_{ST}$ if populations have diverged for a sufficiently long time in the case of ancient isolation (Hardy *et al.*, 2003). To visualise the relationship among samples, MDS analysis based on $F_{ST}$ and $R_{ST}$ values were carried out in the R environment, and population structure of individual genotypes was visualized by correspondence analysis (CA) using GENETIX 4.05 (Belkhir *et al.*, 1996).

We tested concordance of pairwise genetic distances using a Simple Mantel test among redfish collections from available areas, based on mtDNA $\Phi_{ST}/F_{ST}$ vs. microsatellite $F_{ST}$. To further illustrate the distribution pattern of horizontal and vertical genetic variation, partial mantel tests were calculated between matrices of genetic distances (using mtDNA and microsatellite markers) and depth (m) while removing the influence of geographic distance (km), and vice versa, using the Vegan (Oksanen *et al.*, 2011) package in R, with statistical significance evaluated with 9,999 permutations.

We then used Approximate Bayesian Computation (ABC), implemented in DIYABC v2.0 (Cornuet *et al.*, 2008, 2014), to assess the evolutionary relationships among *S. mentella* lineages in the North Atlantic. We compared six simple scenarios (see results), to identify the most likely evolutionary events, and the effective population sizes and times of divergence implicated. The ABC approach is used to provide inference about the posterior distribution of the model parameters in order to establish how the summary statistics compare between simulated and observed datasets (Cornuet *et al.*, 2008). We used default priors and generated simulated datasets combining both mtDNA and microsatellite information, with a total of 28 summary statistics, the calculation of which is implemented in DIYABC; for microsatellites: mean number of alleles, genic diversity, size variance and Garza & Williamson's M-ratio per population and $F_{ST}$ pairwise divergence; for mtDNA: $F_{ST}$ pairwise divergence only. Additional summary statistics generated for

mtDNA failed to provide satisfactory statistical fit between the observed and simulated datasets (Data not shown). Runs consisted of six million simulated datasets (one million for each scenario) and evaluated by Principal Components Analysis (PCA), based on the obtained summary statistics of a subset of simulations, to check the threshold distance of parameter estimates between simulated and observed data. The relative posterior probability was then estimated over each scenario via a logistic regression on 1% of simulated datasets closest to the observed dataset. The posterior distribution of parameters for the summary statistics associated with the retained dataset was then estimated through locally weighted linear regression.

## 2.4 Results

### 2.4.1 Mitochondrial DNA lineages

MtDNA sequence data of 261 *S. mentella* individuals from 16 collections stretching across the North Atlantic produced a 444 bp fragment alignment with a total of 44 polymorphic sites, 25 of which were parsimony informative (accession numbers: KP988027-KP988288). These polymorphisms defined 56 haplotypes with overall haplotype diversity and mean nucleotide diversity of $h$=0.897±0.009 and $\pi$=0.005±0.0002 respectively (Table 2.1). The haplotypes were organised into two main divergent haplogroups differentiated by a mean net sequence divergence percentage ($D_a$) of 2.45 ± 0.47 (Figure 2.3a). These haplogroups "clade A" and "clade B" correspond to the "shallow" and "deep" groups previously described by Shum *et al.* (2014 and Figure S7.1, supplementary material).

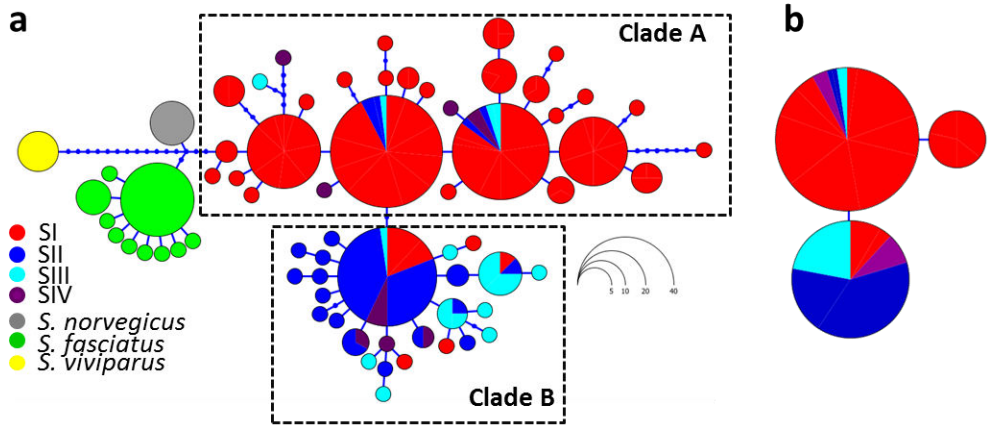



**Figure 2.3** Haplotype genealogy of North Atlantic *Sebastes* sequences yielded using HaplotypeViewer, using maximum likelihood. (a) 68 mitochondrial control region *Sebastes mentella* haplotypes are ordered into clade A and clade B; (b) three rhodopsin haplotypes ordered as shallow and deep groups. Haplotypes are coloured according to SAMOVA groupings (SI-SIV, See Table 2.3), and the size of each circle is proportional to the frequency of haplotypes. The connecting lines reflect the number of mutations between them.

Among all samples, 91.7% of redfish caught above 500m (59.5%) were classified as clade A and 8.3% as clade B haplotypes, whereas among redfish caught below 500m (40.5%) 34% belonged to clade A and 66% to clade B haplotypes. Clade A included 68.3% of the total haplotypes found in 16 collections with 12 shared haplotypes (43%) and 16 singletons (57%). Clade B haplotypes (31.7%) formed a far more circumscribed starburst pattern, with haplotypes all found in 8 of the 16 locations with 6 shared haplotypes (21.1%) and 22 singletons (79%), suggestive of a recent expansion. The clades generally show an association to both location ($\chi^2$= 180.75, p<0.0001, df=15) and depth ($\chi^2$=61.76, p<0.0001, df=1), with only a few examples of shared haplotypes between adjacent groups (Figure 2.3a). Within the Irminger Sea, clade A haplotypes dominated waters above 500m (75%) in the northeast, west and southwest, while clade B was prevalent at depths below 500m (90%) in the northeast and west (Figure 2.2).

Overall, mtDNA estimates of genetic differentiation for all comparisons involving the Northeast Irminger Sea, west Faroe and Norwegian Sea samples were significant ($\Phi_{ST}$=0.055-0.423 and $F_{ST}$=0.100-0.635), while the remaining samples show no significant genetic structure and low heterogeneity (Table 2.2). The MDS plot (Figure 2.4a, b) based on mtDNA ($\Phi_{ST}$ & $F_{ST}$) pairwise genetic distances show consistent subdivision of the shallow-pelagic clade A from the deep-pelagic clade B groups along axis 1. Axis 2 shows a

separation between the deep-pelagic west Faroe from the NE Irminger Sea groups (>500m).

The results of SAMOVA indicated significant population genetic structure for each number of *k* groups assumed. $F_{CT}$ in the SAMOVA analysis showed increasing values of *k* up to 4 and began to progressively decline after that, while $F_{SC}$ drops significantly at *k*=4, which resulted in the maximum variance between the two indices at this point. Further increases in the number of *k* led to a dissolution in group structure, where groups with the larger proportion of private haplotypes were singled out. Thus, we chose four subpopulations as the most parsimonious partition scheme (Figure 2.5). At *k*=2, SAMOVA recovered a deep mitochondrial split between the combined "deep-pelagic" groups from northeast Irminger Sea (>500m) and west Faroe Island, and everything else. At *k*=3, the northeast Irminger Sea (>500m) group and the west Faroe group became separated. At *k*=4, redfish from the majority of clade A caught above 500m depth formed subpopulation SI (see Table 2.3), the northeast Irminger Sea and west Faroe groups (deeper than 500m), representing the core of clade B, formed subpopulation SII and SIII, respectively; finally, some redfish from the northeast Irminger Sea (<500m; n=11, clade A: 45%, clade B: 55%) formed subpopulation SIV. A hierarchical AMOVA analysis showed the molecular variance attributed to variance among groups to account for 37.54% ($F_{CT}$= 0.375, $p$<0.001), while differentiation between collections within the same group ($F_{SC}$) was found not to be significant ($F_{SC}$= 0.013, $p$= 0.123). The analysis using SAShA revealed the average distance between co-occurring haplotypes was 951.9km (geographic distance) and 178.3m (depth), with evidence that haplotypes were geographically restricted as the observed distribution (OM) was significantly different from the expected (EM) under the assumption of panmixia for geographic distance (D$g$=80, $p$=0.018) and depth (D$g$=11.6, $p$=0.039).

Demographic analysis on a significantly expanded data set showed similar results to Shum *et al.*, (2014): unimodal distributions with negative and significant Tajima's *D* and Fu's *F* values for clade A (*D*= -1.67, $p$<0.01; *F*= -20.57, $p$<0.001) and clade B (*D*= -2.14, $p$< 0.001; *F*= -27.09, $p$<0.001), providing support for recent population expansion (Figure S7.2). We find that the "deep" and "shallow" lineages split around 22,000±9,000 years ago. The coalescent-based Bayesian skyline plot (BSP) provided details on how mtDNA diversity changed through time, indicating a rapid population growth for clade A in the past 10-12

kyr before present (BP), with a very recent stall/reduction in effective population size (Figure 2.6a). Clade B haplotypes demonstrate a signal of a sustained period of population increase dating back over 15-25 kry BP (Figure 2.6b).

### 2.4.2 Comparative nuclear and mitochondrial data

Rhodopsin sequence data from 160 *S. mentella* individuals (accession numbers: KR818563–KR818700) produced a 722 bp fragment alignment with a total of 2 polymorphic sites at positions 208 and 228. These polymorphic sites defined 3 haplotypes structured into well-defined shallow and deep rhodopsin clades (Figure 2.3b). The Single Nucleotide Polymorphism (SNP) T228 was only variable among the "shallow" group; however one SNP found at position 208 was fixed for alternative alleles between shallow (G) and deep (A) fish, strengthening Shum *et al.*'s (2014) findings. The distribution of the SNP genotypes at position 208 show a strong association to mtDNA clade A and clade B haplogroups for the shallow ($\chi^2$= 675.17, p<0.0001) and deep ($\chi^2$= 37.28, p<0.0001) collections respectively, with 91% of shallow "G" rhodopsin variants also belonging to mtDNA clade A, and 64% of deep "A" rhodopsin genotypes being assigned to mtDNA clade B. Microsatellite pairwise FST values between the nine geographically defined regions ranged from -0.003 to 0.040 (Table S7.1) and were significant for most combinations that included the Irminger sea (>500m) and Norwegian International Water collections, after FDR adjustments. Overall, 13 of the 36 pairwise comparisons among historical samples resulted in a significant phylogeographic signal (RST significantly larger than FST, see Table S7.1, Figure 2.4c, d). The Correspondence Analysis clearly segregated redfish into three main groups. Axis1 separated the "deep-pelagic" clade B groups ($IrNE_{DP}06$ & $FW1_{DP}07/FW2_{DP}07$) from "shallow-pelagic" clade A groups ($IrSW_{SP}06$, $FE1_{SP}07/FE2_{SP}07$, $NS_{SP}07$ & $NIW1_{SP}06/NIW2_{SP}06$), and Axis 2 separated the west Faroe ($FW1_{DP}07/FW2_{DP}07$) from the east Faroe ($FE2_{SP}07$) and northeast Irminger Sea (>500m: $IrNE_{DP}06$) groups (Figure 2.4e).

**Table 2.2** MtDNA estimates of pairwise genetic differentiation among 16 *S. mentella* collections.

| | Irminger Sea | | | | | | | | | Faroe West | | Faroe east | | NS | NIW | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $SW_{DP}13$ | $SW_{SP}13$ | $SW_{SP}06$ | $W_{SP}13$ | $W_{DP}13$ | $W_{INT}13$ | $NE_{DP}13$ | $NE_{SP}13$ | $NE_{DP}06$ | $FW1_{DP}07$ | $FW2_{DP}07$ | $FE1_{SP}07$ | $FE2_{SP}07$ | $NS_{SP}07$ | $NIW1_{SP}06$ | $NIW2_{SP}06$ |
| **$SW_{DP}13$** | - | 0.033 | -0.040 | -0.033 | 0.025 | 0.029 | **0.530***** | **0.178** | **0.615***** | **0.333**** | **0.362***** | 0.001 | 0.038 | 0.080 | 0.013 | -0.077 |
| **$SW_{SP}13$** | 0.033 | - | -0.003 | -0.011 | 0.052 | 0.043 | **0.517***** | **0.219***** | **0.560***** | **0.406***** | **0.408***** | 0.055 | 0.027 | 0.139 | -0.030 | 0.008 |
| **$SW_{SP}06$** | -0.007 | -0.020 | - | -0.030 | 0.040 | 0.027 | **0.515***** | **0.213***** | **0.566***** | **0.383***** | **0.394***** | 0.037 | 0.017 | **0.116** | -0.014 | -0.032 |
| **$W_{SP}13$** | -0.008 | -0.024 | -0.029 | - | 0.022 | 0.011 | **0.474***** | **0.173**** | **0.524***** | **0.329***** | **0.340***** | 0.013 | 0.015 | **0.097** | -0.024 | -0.033 |
| **$W_{DP}13$** | -0.010 | 0.019 | 0.001 | -0.001 | - | -0.039 | **0.302***** | 0.035 | **0.345***** | **0.179**** | **0.187**** | 0.011 | 0.039 | **0.092** | 0.022 | 0.015 |
| **$W_{INT}13$** | 0.006 | 0.005 | -0.011 | -0.010 | -0.035 | - | **0.354***** | 0.049 | **0.409***** | **0.202**** | **0.210**** | -0.012 | 0.057 | **0.134** | 0.006 | 0.007 |
| **$NE_{DP}13$** | **0.179**** | **0.159***** | **0.161***** | **0.156***** | **0.066** | 0.069 | - | **0.111** | 0.001 | **0.160**** | **0.146**** | **0.429***** | **0.538***** | **0.529***** | **0.529***** | **0.517***** |
| **$NE_{SP}13$** | **0.106** | 0.049 | 0.055 | 0.053 | 0.015 | -0.003 | 0.025 | - | **0.134** | 0.064 | 0.076 | **0.096** | **0.167** | **0.198** | **0.157** | **0.172** |
| **$NE_{DP}06$** | **0.304***** | **0.250***** | **0.257***** | **0.254***** | **0.143**** | **0.152**** | 0.003 | 0.082 | - | **0.203**** | **0.192***** | **0.494***** | **0.635***** | **0.628***** | **0.608***** | **0.599***** |
| **$FW1_{DP}07$** | **0.112** | **0.074** | **0.077** | **0.073** | **0.061** | 0.048 | **0.118** | 0.028 | **0.209**** | - | -0.080 | **0.214**** | **0.437***** | **0.456***** | **0.371***** | **0.346***** |
| **$FW2_{DP}07$** | 0.074 | **0.058** | **0.058** | 0.053 | **0.062** | **0.055** | **0.154**** | 0.057 | **0.260***** | -0.049 | - | **0.230**** | **0.445***** | **0.464***** | **0.379***** | **0.363***** |
| **$FE1_{SP}07$** | -0.049 | 0.024 | 0.012 | -0.001 | 0.018 | 0.032 | **0.181**** | **0.100** | **0.302***** | 0.079 | 0.041 | - | 0.073 | **0.133** | -0.020 | -0.020 |
| **$FE2_{SP}07$** | 0.004 | 0.032 | 0.020 | 0.016 | 0.056 | 0.075 | **0.243***** | **0.164**** | **0.369***** | **0.188**** | **0.137** | -0.029 | - | -0.014 | 0.005 | 0.004 |
| **$NS_{SP}07$** | 0.048 | **0.145**** | **0.136**** | **0.123** | **0.136**** | **0.182**** | **0.298***** | **0.265***** | **0.423***** | **0.279**** | **0.216**** | 0.010 | -0.017 | - | 0.128 | 0.071 |
| **$NIW1_{SP}06$** | 0.054 | -0.023 | -0.019 | -0.017 | 0.035 | 0.021 | **0.195***** | 0.067 | **0.313***** | **0.102** | 0.076 | 0.007 | -0.016 | 0.145 | - | -0.034 |
| **$NIW2_{SP}06$** | -0.083 | 0.010 | -0.017 | -0.024 | -0.005 | 0.002 | **0.188***** | **0.108**** | **0.309***** | **0.121**** | **0.082** | -0.049 | -0.047 | 0.031 | 0.010 | - |

$\Phi_{ST}$ below diagonal and $F_{ST}$ above diagonal at 0.05 *p*-value significance following False discovery rate (FDR) correction. <0.05 BOLD; <0.01 **; <0.001 ***

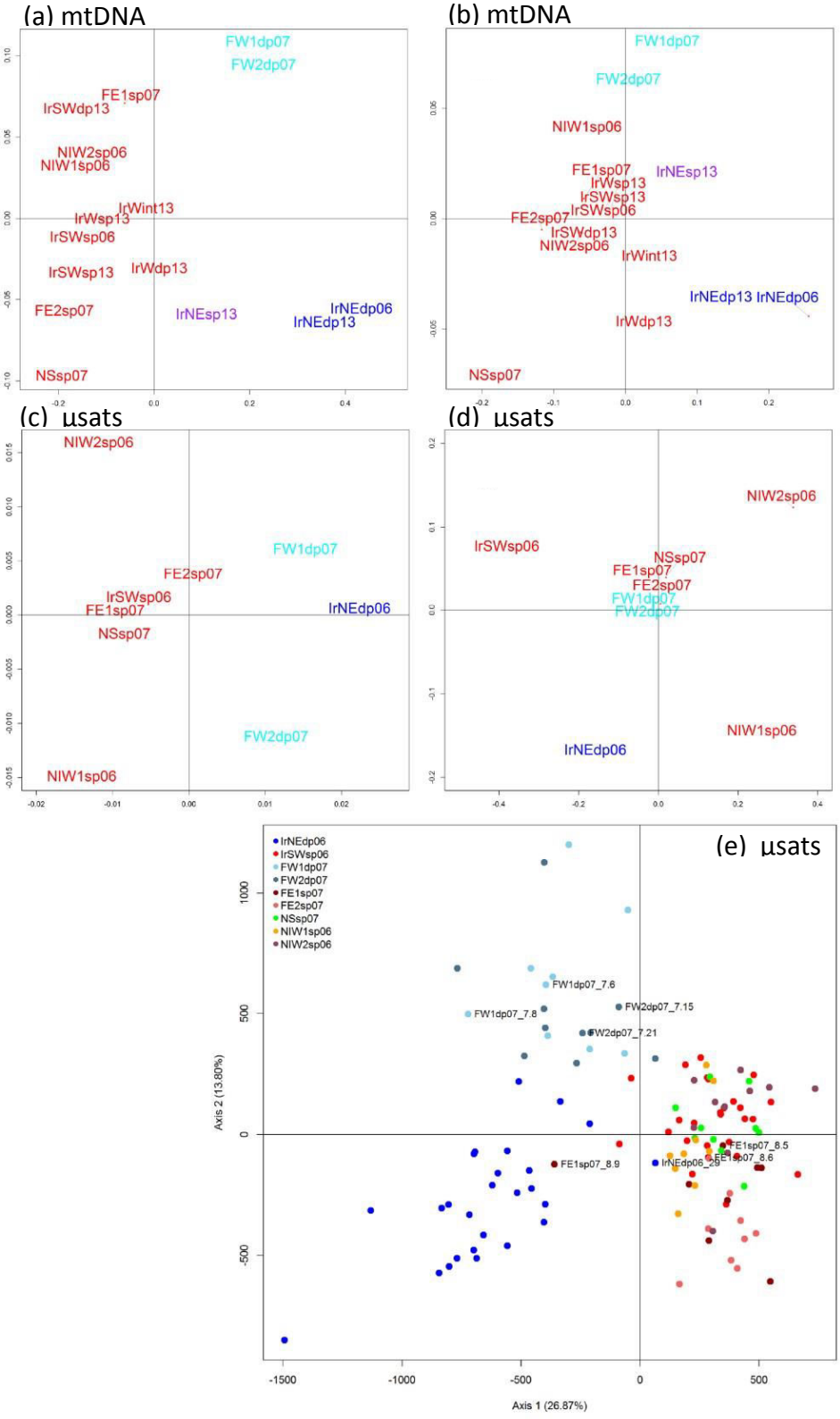



**Figure 2.4** Multidimensional Scaling Plot (MDS) of *Sebastes* mtDNA differentiation from 16 collections (a) FST, (b) ΦST; microsatellite variation from 9 collections (c) FST and (d) RST pairwise genetic distances across the North Atlantic (Location codes (see Table 2.1) coloured according to SAMOVA groupings); (e) correspondence analyses of microsatellite *S. mentella* samples from 9 locations.

**Table 2.3** mtDNA analysis of molecular variance for SAMOVA groupings for *S. mentella* from 16 localities in the North Atlantic. SP: shallow-pelagic (<500m), DP: deep-pelagic(>500m), INT: intermediate (between 400-600m)

| Groups | Group names | Group index | | | | |
|---|---|---|---|---|---|---|
| | | Tajima's D | Fu's *F*s | Source of variation | Percentage of variation | Fixation indices |
| **1. $IrW_{Int}13+IrW_{DP}13+IrW_{SP}13+IrSW_{DP}13+IrSW_{SP}13+IrSW_{SP}06+FE1_{SP}07+FE2_{SP}07+NS_{SP}07NIW1_{SP}06+NIW2_{SP}06$** | SI | **-1.82**** | **-22.62***** | Among groups | 37.54 | $F_{CT}$= **0.375***** |
| **2. $IrNE_{DP}06+IrNE_{DP}13$** | SII | **-2.02**** | **-16.45***** | Among populations within groups | 0.84 | $F_{SC}$ =**0.013***** |
| **3. $FW1_{DP}07+FW2_{DP}07$** | SIII | -0.47 | **-5.17**** | Within populations | 61.62 | $F_{ST}$= **0.383***** |
| **4. $IrNE_{SP}13$** | SIV | 1.43 | 2.47 | | | |

<0.05 BOLD; <0.01 **; <0.001 ***

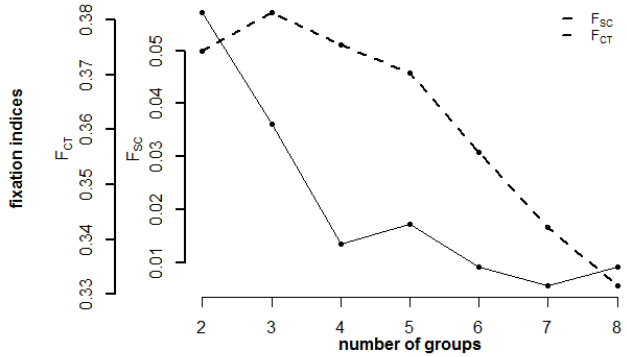



**Figure 2.5** Spatial analysis of molecular variance. Fixation indices obtained for eight different aggregations of samples tested using SAMOVA. Y-axis show $F_{CT}$ and $F_{SC}$ fixation indices and x-axis show number of groups (k). $F_{CT}$: variation among groups of populations and $F_{SC}$: variation among populations within groups.

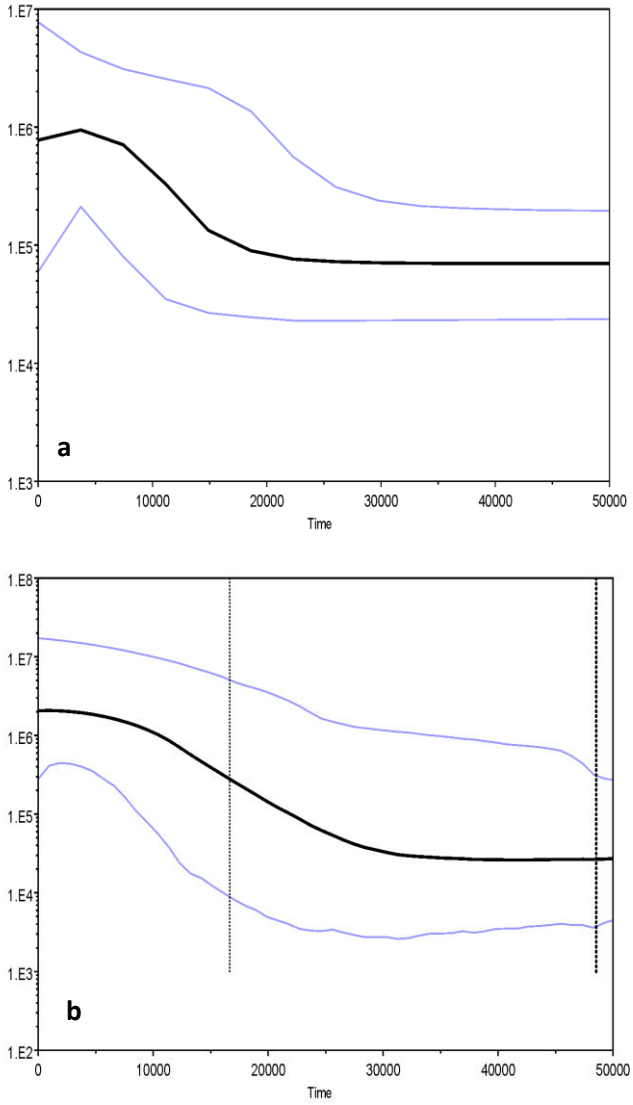



**Figure 2.6** Bayesian Skyline Plots (BSPs) showing the temporal changes in mtDNA diversity. The x-axis is in years and y-axis show the genetic diversity expressed as the product of effective female population size. (a) clade A, (b) clade B.

North Atlantic structuring of redfish based on microsatellite markers and results from mtDNA (SAMOVA) consistently divided the "deep-pelagic" groups (Irminger Sea (>500m) & west Faroe) into two subgroups. We found concordance between pairwise genetic distances among redfish in 9 localities based on mitochondrial $\Phi_{ST}/F_{ST}$ vs. microsatellite $F_{ST}$, showing a significant positive correlation of the measures of genetic differentiation (Figure S7.3, tested with 9999 permutations). We found no significant signal of isolation-by-distance (r=0.12, $p$=0.24) or depth (r=0.12, $p$= 0.25) based on microsatellite $F_{ST}$, using a partial mantel test where the depth/geographic distance matrix was held constant. For mtDNA, partial mantel tests showed a significant correlation between genetic distance ($F_{ST}$) and depth ($r$=0.329, $p$=0.019), but not for geographic distance ($r$=0.009, $p$=0.407).

Given the detection of geographically disjunct "deep" populations in both the Irminger and Faroese Seas, we tested six simple historical demographic scenarios involving pairs of "shallow" and "deep" populations from these areas (bearing in mind the relatively consistent homogeneity on the "shallow" populations all across the NA, Figure 2.2, Figure 2.5). The first scenario involved an original split between the "deep" and "shallow" lineages followed by subsequent split between the Irminger and Faroe Seas within both deep and shallow lineages, leading to the current geographical disjunction. The second scenario involved a parallel, independent origin of deep sea groups from their shallow ancestors respectively in the Irminger and Faroese Seas. The third differed from the second scenario with an independent origin of the shallow groups from their deep sea ancestors. The fourth, fifth and sixth scenarios involved a simplistic treelike bifurcations occurring from time *t*a to *t*1 in the past. Estimations of the posterior probability based on both direct and logistic regression for each scenario provided unambiguous support for the scenario S1, with a probability of 91% and a 95% CI of 90–92, not overlapping with any other scenarios (Figure 2.7). Given a generation time between 12-20 years (Stransky *et al*., 2005a, b), the scenario assumes that the "deep-pelagic" and "shallow-pelagic" redfish split ~4,500-7,500 years ago (*t*a median = 374 generations, 95% CI: 93.4-4190). The subsequent split between the "deep-pelagic" Irminger Sea ($IrNE_{DP}06$: $N_e$ = 8570, 95% CI: 5840-9900) and west Faroe ($FW_{DP}06$: $N_e$ = 7040, 95% CI: 3160-9800) share similar divergence times as the split between the "shallow-pelagic" Irminger Sea ($IrSW_{SP}06$: $N_e$ = 6610, 95% CI: 3310-9640) and east Faroe ($FE_{SP}06$: $N_e$ = 5570, 95% CI: 2020-9510), estimated at approximately ~1,200-2,000 years ago (t1 median = 98.3 generations, 95% CI: 27.9-281).

## 2.5 Discussion

Based on greatly expanded geographical screening, sample sizes and analytical toolkit, this study lends strong support to the existence of at least two highly distinct evolutionary units of oceanic *Sebastes mentella* in the North Atlantic (NA). Nuclear patterns are mirrored and strengthened by mitochondrial evidence, and, perhaps more strikingly, the pattern of divergence is strongly associated with habitat depth (more so than with geographical distance). Coalescent patterns suggest that the emergence of depth-related structure may have originated in the late Pleistocene, with the two lineages segregating after the

recolonization of the Irminger Sea, as the ice sheets retreated at the end of the last glaciation.

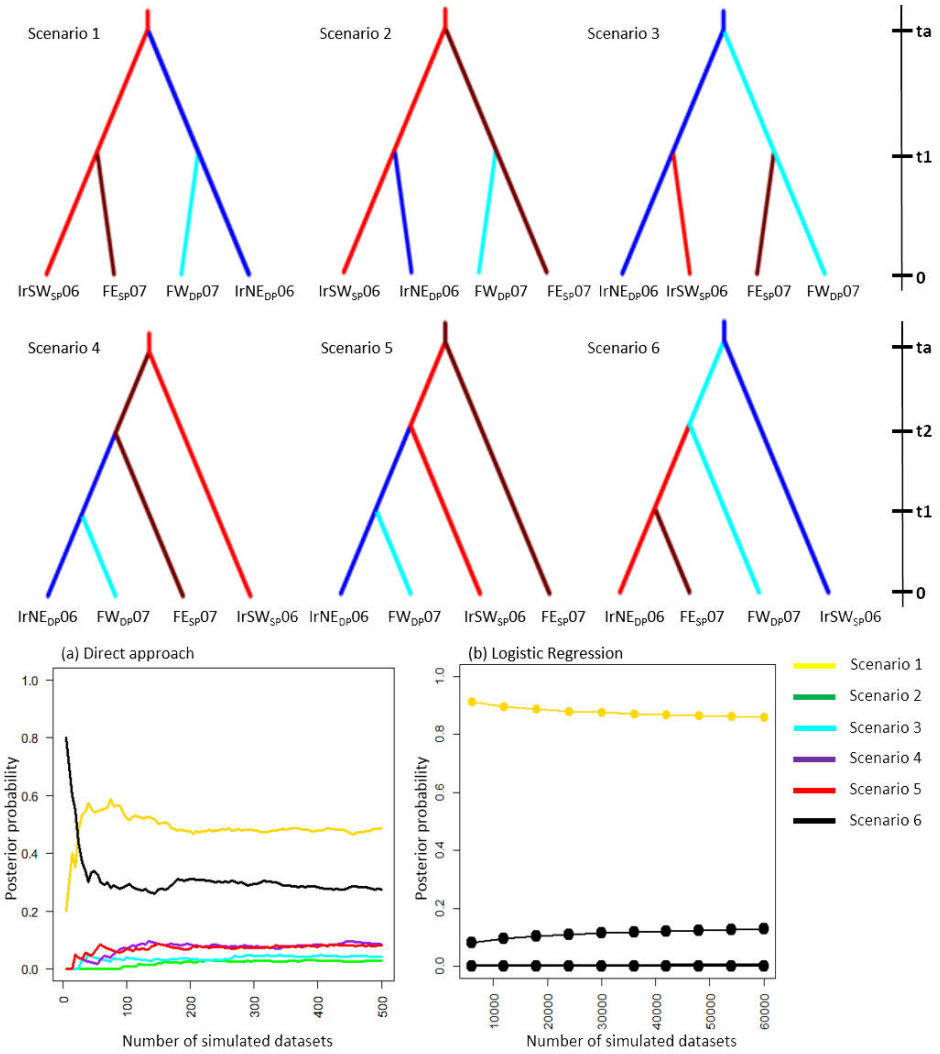



**Figure 2.7** Alternative scenarios tested using Approximate Bayesian Computation approach (DIYABC) based on combined microsatellite and mtDNA data for *S. mentella*. Top panel illustrate the six scenarios tested. Bottom panel indicate the relative likelihoods of the six scenarios compared by (a) direct approach and (b) logistic regression on the 1% (60,000) and 0.008% (500) of the closest simulated data sets, respectively. On the y-axis, the posterior probabilities, and on the x-axis the number of simulations used to calculate it (1% of total simulations). The DIYABC analysis illustrate scenario 1 as the best-supported scenario. See Table 2.1 for corresponding location codes.

### 2.5.1 Population structure and connectivity

Analyses of mtDNA sequences revealed substantial population structure among depth defined habitats between two distinct *S. mentella* mtDNA clades (A: above and B: below 500m; Figure 2.2 and Figure 2.3). Our results expand on recent findings, indicating that the "shallow" and "deep" groups in Shum *et al.* (2014) corresponds to clades A and B

identified in this study (Figure S7.1). We found non-random association of haplotypes in the NA, with clade A commonly found above 500m southwest and west of the Irminger Sea, east of the Faroes, and Norwegian waters, whereas clade B haplotypes are more localised to the central NA, below 500m, northeast of the Irminger Sea and west of the Faroes (Figure 2.2). The overall pattern of mtDNA variation is associated with isolation-by-depth, mirrored by rhodopsin SNP distribution, similarly detected using microsatellites (Stefánsson *et al.*, 2009b), and to some degree also supported by subtle phenotypic variation (Magnússon & Magnússon, 1995; Stransky 2005; Stefánsson *et al.*, 2009a,b; Cadrin *et al.,* 2010). MtDNA and nuclear data consistently identified the main genetic partitioning – shallow vs deep – as well as detecting additional population subdivision, between the "deep-pelagic" Irminger Sea and west Faroe groups. According to the SAMOVA and MDS analysis, redfish was further divided into four genetically distinct subclades; (SI) southwest and west Irminger Sea/ east Faroe/ Norwegian shelf and Norwegian International waters; (SII) northeast Irminger Sea (>500m); (SIII) west Faroe; and (SIV) northeast Irminger Sea (<500m). However, it is worth mentioning the potential issue of small sample size for some collections as is the case of the final SAMOVA group (SIV), which likely represents an area of contact between haplogroups and may not reflect a "true" distinct biological unit but a mere artefact due to low sample size. Support for this group structure is mirrored by close inspection at nuclear markers. The rhodopsin SNP shows a strong association with the mtDNA clade A and clade B haplotypes, suggesting that evolutionary independence between shallow and deep lineages is unambiguous. The two deep-pelagic groups were not statistically distinguishable at microsatellite loci, even using greater sample sizes (Stefánsson *et al.* 2009b). Yet, analysed at mtDNA, they showed significant genetic structure (Figure 2.2a & Figure 2.4). Strikingly, while the average mtDNA-based $F_{ST}$ variance among "shallow" aggregations is non-significant ($F_{ST}$=0.020, $p$=0.074), suggesting significant large scale connectivity, the two "deep" groups, in the Irminger Sea and Western Faroes, appear significantly differentiated ($F_{ST}$=0.176, $p$>0.001), with only three haplotypes shared, indicating that habitat segregation may have shaped diverging behaviours and life-history adaptations in these two lineages. Such a segregation of the mitochondrial matrilineage, between deep Irminger and Faroese fish, alongside a more blurred boundary showed by microsatellites, raises the possibility that female *S. mentella* in the deeper sea layer may exhibit some degree of female philopatry/residency (Petit & Excoffier, 2009). The combined use of mitochondrial and nuclear information on a few selected individuals can also shed some further light on the

biology and behaviour of these organisms. For example, one fish caught above 500m from east of the Faroes ($FE1_{SP}07\_8.9$) exhibits a clade B haplotype, typical of the "deep" layer, and its microsatellite genotype falls within the "deep" group. This highlights the ability of redfish to exhibit a degree of pelagic-demersal mixing between clades from shallow and deep environments (Planque *et al*., 2013). The inception of this behaviour may originate during the establishment of juvenile "site-fidelity" during the early stages of redfish development. In the Irminger and Norwegian Seas, spawning female redfish are distributed along the larval extrusion areas of the Reykjanes ridge and the Norwegian continental shelf between March and May (Magnússon & Magnússon, 1995; Drevetnyak & Nedreaas, 2009). The pelagic-fry drift to nursery areas found along the coast of east/west of Greenland and the Barents Sea where juveniles settle to the bottom until they mature and migrate to distributed populations throughout the North Atlantic. Yet, the complex behaviour, areas of copulation and seasonal migration patterns of this species across shallow- and deep-pelagic habitats remain poorly understood (Planque *et al*., 2013). Hence this pattern suggests the occurrence of short-time migration from deep-pelagic individuals, supporting the notion of vagrants as individuals move to shallow and deeper waters during their life-cycle (Shum *et al*., 2014).

Several examples from the east/west Faroe groups caught in both shallow and deep layers display discordant molecular markers: two samples ($FE1_{SP}07\_8.5$ & $FE1_{SP}07\_8.6$) caught approximately at 400m display "deep" haplotypes but a microsatellite genotype typical of the shallow layer. Similarly, five samples caught in deep waters ($FW1_{DP}07\_7.6$, $FW1_{DP}07\_7.8$, $FW2_{DP}07\_7.15$, $FW2_{DP}07\_7.21$ & $IrNE_{DP}06\_29$) exhibited clade A haplotypes but possess multilocus microsatellite genotypes falling within the "deep" group. The nuclear-coding rhodopsin gene is a powerful tool in the identification of closely related fish species (Rehbein 2013; Shum *et al*., 2014; Pampoulie *et al*., 2015) and here we found that only 7 out of 160 specimens (4%) analysed at both mtDNA and rhodopsin yielded ambiguous assignment to their lineage of origin (Table S7.2). Interestingly, two of these individuals possessed double peaks or heterozygous indels in the rhodopsin chromatograms (Figure S7.4; Sousa-Santos *et al*., 2005), which, along with the observed mtDNA/microsatellite mismatches, is best explained by the occasional occurrence of introgressive hybridization between the two genetically distinguishable groups. Roques *et al*. (2002) indicated significant introgressive hybridization between *S. mentella* and *S. fasciatus* in the Gulf of St. Lawrence and surrounding waters; Pampoulie & Danielsdottir (2008) also detected signatures of hybridisation between *S. mentella* and *S.*

*norvegicus*. Thus, it is reasonable to expect that intraspecific mating between a deep-pelagic clade B male and shallow-pelagic clade A females (and *vice versa*) would have produced the genotypic ambiguities mentioned above. The shallow and deep-pelagic groups appear to occupy their preferential depth range in near-sympatry. However, spatial overlap is more apparent to the west and northeast of the Irminger Sea and in the Faroese waters, thus the opportunity for introgressive hybridization may derive from hybrid zones as a result of secondary contact where *S. mentella* are within their preferred geographic and vertical cruising limits. Pampoulie and Daníelsdóttir (2008) have reported significant levels of hybridization in the NA among redfish which may indicate that different *S. mentella* lineages were allopatric before secondarily coming into contact to form their current sympatric distribution (Cadrin *et al*., 2010). Despite the potential for interbreeding, however, the shallow and deep types maintain a striking overall integrity, which bears resemblance to some notable cases of parapatric speciation (Allender *et al*., 2003; Berner *et al*., 2009; Nosil, 2012).

### 2.5.2 Historical reconstruction

The glaciation and interglacial periods of the Pleistocene have been known to considerably shape the current distribution and connectivity among contemporary populations (Finnegan *et al*., 2013; Bigg *et al*., 2015). Analysis of mtDNA haplotypes alone suggest that the shallow- and deep-pelagic clades diverged over 22,000±9,000 years ago, in correspondence with the Last Glacial Maximum (LGM). The Bayesian Skyline Plot (BSP) provides information on how mtDNA diversity changed through time, back to the most recent common ancestor. The BSP analysis indicate a gradual signature of expansion 15,000-25,000 years ago for the deep-pelagic clade, and a greater, steady increase in population growth - compared to the shallow-pelagic clade – which shows a much sharper rapid expansion around 10,000-15,000 years ago. The BSPs show contrasting signatures of population growth suggestive of postglacial influence. The expanding sea ice may have driven *S. mentella* to southerly latitudes as far as the Grand Banks to deeper waters during the LGM, as global sea level fell by 120-135m and intense calving in the northern hemisphere ice sheets 18-15 kyr BP resulted in massive icebergs advancing into the North Atlantic 40˚N (Grousset *et al*., 1993). Advancing ice would have caused low productivity in areas of ice cover which would have resulted in reduced photosynthesis and food availability. As the sea ice retreated, characterized by intense warming and rising sea

levels 14,600-13,800 years BP, *S. mentella* had the opportunity to advance in a northerly direction to glacial refugia south of Iceland before sea ice readvanced during the Younger Dryas (YD, 13-11kyrBP) which forced the northern hemisphere into near-glacial conditions (Crucifix & Berger, 2002). The onset of the YD may have favoured allopatric conditions between the two clades as the deep-pelagic group occupied deeper refuges, while the shallow-pelagic group rapidly spread with the retreating sea ice and rising sea level following the YD into the Norwegian and Barents Seas and the Irminger Sea ($R_{ST}$>$F_{ST}$; Table S7.1). The microsatellite $R_{ST}$ values provide insights on the relative divergence of *S. mentella*. The eastern NA collections ($FE_{SP}07$, $NS_{SP}06$ & $NIW_{SP}06$) present a significant phylogeographic signal with respect to the Irminger Sea collections ($IrSW_{SP}06$ & $IrNE_{DP}06$), suggesting that post-glacial range shifts and secondary contact events have played a significant role in shaping *S. mentella* spatial structure.

Overall, SAMOVA, MDS and CA analyses revealed the shallow-pelagic group show strong homogeneity across the NA, whereas the deep-pelagic groups form rather distinct "pockets", suggesting their evolutionary distinction also involves different life-histories, as shallow-pelagic redfish appear more prone to connectivity, whereas the deep-pelagic redfish appear to be more strictly associated with their local habitats, probably migrating less. We tested which demographic scenario could best explain the genetic patterns among pelagic pairs of shallow and deep populations in the Irminger Sea and West of the Faroes, using an ABC approach. Estimates for the posterior probability for each scenario provided robust support for scenario 1 (Figure 2.7). This indicates that the shallow- and deep-pelagic gene pools split approximately 5-7 kyr BP, after the postglacial recolonization of the Irminger Sea. Furthermore, a subsequent split between the deep-pelagic groups emerged approximately 2 kyr BP. Their distributions are found within close proximity of bathymetric features (see Figure 2.2) along the West of the Reykjanes ridge and the Faroe-Shetland ridge known for complex oceanic conditions at depth (Pedchenko, 2005; Olsen *et al.*, 2008). This may act as bathymetric forcing of ocean currents as reported for North Atlantic tusk (Knutsen *et al.*, 2009). The combination of these physical obstacles and the presumably complex and largely unexplored reproductive behaviour of *S. mentella* may have acted in concert initiating the retention and establishment of reproductive isolation by depth (Shum *et al.*, 2014).

## 2.6 Conclusions

Our study provided the first extensive depth associated sampling and phylogeographic reconstruction of the North Atlantic oceanic beaked redfish, *Sebastes mentella*. This species shows a consistent and significant distinction of at least two evolutionary lineages: one widely-distributed "shallow" type, which partially overlaps with local populations of a more sedentary "deep" type. While evidence exists for localised interbreeding between these two lineages, the rate does not appear notably greater than similar introgressive processes still occurring among the four *Sebastes* species in the North Atlantic. We find that mtDNA reflects redfish evolutionary history throughout the late Pleistocene, whereas the integration with microsatellite data serves to better reflect postglacial divergence and the patterns of contemporary gene flow among populations, including the potential detection of sex-biased dispersal. The shallow-pelagic clade show strong homogeneity across the NA, while the deep-pelagic clade is restricted to the central North Atlantic, segregated by complex oceanic conditions shaped by bathymetric features. Population independence was largely upheld by both nuclear and mitochondrial markers and it is likely that depth-associated adaptive processes are at play (Shum *et al.*, 2014) to counteract the homogenising effects of gene flow. Given the paucity of well-characterised marine biological systems undergoing diversification processes consistent with speciation, North Atlantic *S. mentella*, should now represent a valuable subject to investigate genomic correlates and mechanisms for the maintenance of bathymetry-associated lineage sorting and, potentially, speciation.

From a practical standpoint, the present results will have cascading impacts on the assessment and management of these commercially valuable fish stocks, which remain hotly debated. Our results are consistent with two genetically distinguishable putative groups separated by depth and following different evolutionary trajectories: this should form the basis to recognise them as distinct evolutionarily-significant units. Given the circumscribed local distribution of the "deep" populations in the Irminger and Western Faroese seas, tailored management appears required, if we are to avert the permanent loss of a unique biodiversity component.

# Chapter III

# The Visual Pigment in *Sebastes mentella* and its Significance for Habitat Depth

## 3.1 Abstract

The visual pigments of fish are thought to be adapted to the variable spectral properties of their surrounding light environments. Depth arguably plays an important role in shaping adaptation to the dim-light gradients of oceanic habitats, and evidence exists that shows rhodopsin is positively selected to different light conditions. The beaked Atlantic redfish, *Sebastes mentella*, exhibits primarily pelagic, open ocean behaviour, occurring in assemblages down to 1000m depth, and experience astonishing degree of variation in light levels. Here we present the first attempt to empirically link rhodopsin genotype variation with retinal spectral absorption between fish dwelling at different habitat depths in the North Atlantic. Rhodopsin variation in *S. mentella* shows a key amino acid change at site 119 that strongly discriminates the shallow- and deep-pelagic *S. mentella* groups, but is not involved in spectral tuning. Spectral analyses of *S. mentella* visual pigments reveal a single rod-like pigment with average maximal absorbance, $\lambda_{max}$, of 491.8±0.7nm for the shallow-pelagic group (<500m) and 491.1±0.6nm for the deep-pelagic group. Despite the failure to capture an example of adaptive tuning in *Sebastes* by depth, interesting inferences lead to the generation of new hypotheses that require further attention and provides scope for new research avenues to study visual ecology of *S. mentella*.

## 3.2 Introduction

The evolution of visual sensitivity is determined by the spectral range and intensity of available light within a species' visual environment. As a result, animals have different visual pigments in the retina, with spectral sensitivities broadly matching the surrounding light conditions (Lythgoe & Partridge, 1989). Adaptations of visual pigments to the photic environment are achieved by two types of tuning mechanisms: variation in the number of spectral classes of opsin photoreceptors and the variation in the type of chromophore within the photoreceptor outer segment (Bowmaker & Hunt, 2006).

Aquatic animals are found in a range of optical environments due to the physical properties of water, from blue clear oceanic water to turbid brown river water, to light deficient cave environments and the deep-sea. As a result, aquatic animals possess different subsets and combinations of cone photoreceptors (Bailes *et al.*, 2007), to rod-only retinas involved in dim-light vision. This is because deep-dwelling species may have lost functionality of some or all cone opsins given their decreasing importance in darker environments (Yokoyama, 2000; Hunt *et al.*, 2001). In the open ocean, light levels vary considerably with depth, decreasing in intensity and spectral range due to the attenuation of short and long wavelengths. In the mesopelagic zone (200-1000m), very low intensities of sunlight in the blue-green range are able to penetrate down to 1,000m (Warrant & Locket, 2004). This creates a relatively dark environment accompanied by an array of bioluminescent emissions produced by other deep sea animals (between 440-510 nm; Douglas, Mullineaux & Partridge, 2000) that are concentrated in the blue-green part of the visible light spectrum (Widder, 2002, 2010).

The visual sensitivity of marine vertebrates depends on the spectral tuning mechanism of the visual pigment (Yokoyama, 2000), which consists of an integral membrane protein, opsin, and a chromophore (11-*cis*-retinal) attached via a protonated Schiff's base (Yokoyama & Takenaka, 2004). The light sensitivity of the visual pigments is controlled by the interaction between chromophore and opsin and is tuned to a particular wavelength of maximal absorption, $\lambda_{max}$ (Yokoyama, 2000). In animals, visual-system rhodopsins are light-sensitive G-protein-coupled receptors (GPCRs) comprised of an aromatic cluster of amino acids located in the disc membranes of the rod outer segments that activate the specialized heterotrimeric G-protein transducin, and is the pigment involved in dim light vision (Smith, 2010; Yokoyama *et al.*, 2008). Changes in amino acids within the transmembrane domains have been documented to alter the spectral absorbance of rhodopsin in vertebrates (Yokoyama *et al.*, 2008). This raises the hypothesis that shallow-water species may show red-shift in wavelength of maximal absorbance, $\lambda_{max}$, while species that occupy progressively deeper waters are expected to show amino acid substitutions that cause blue shifts in $\lambda_{max}$ (Bowmaker *et al*, 1994). Previous studies have demonstrated cases that predict amino acid positions in the rhodopsin gene that may result in a functional role in different light environments for divergent groups of fishes, namely squirrelfishes (Yokoyama *et al.*, 2008), cichlids (Spady *et al.*, 2005) and *Sebastes* (Sivasundar & Palumbi, 2010). Thus, marine animals living in different photic environments are subject to different evolutionary constraints on their visual system.

The percomorph marine family of rockfishes (Sebastidae) have played a key role in understanding depth-associated divergence and speciation in the ocean (Ingram, 2010; Hyde *et al.*, 2008; Alesandrini & Bernardi, 1999), and evidence exists that rhodopsin gene variation has evolved in response to different depth, and hence, light environments (Sivasundar & Palumbi, 2010). North Atlantic beaked redfish, *Sebastes mentella*, exhibits regional depth-associated substructure (Shum *et al.*, 2014, 2015) with a shallow-pelagic group primarily inhabiting the 200-550m ocean layer over a wide stretch of the North Atlantic, and another, more localised deep-pelagic group dwelling between 550 and 1000m in the Irminger Sea and west of the Faroes. This makes this species a candidate to study phenotypic divergence in response to habitat depth. Nevertheless, the spectral sensitivity of *Sebastes* species remains unknown, so in order to examine this phenotypic trait, a direct measure of the visual pigments is needed to link rhodopsin variation with the spectral absorption for population inhabiting different depths. Here we sampled *S. mentella* at different depths throughout the Irminger Sea, and assessed their visual sensitivity, in an attempt to link genotype and phenotype. We carried out spectral absorbance analysis of 80 *S. mentella* retina along three depth ranges and compared results with rhodopsin SNP variation along the known depth gradient for this species.

## 3.3 Materials and Methods

### 3.3.1 Sample collection

*Sebastes mentella* specimens were collected throughout the Irminger Sea (see section 2.3.1 for further details). Furthermore, the trawl was fitted with a light-tight curtain attached to the central multisampler net, ensuring that the animals inside the codend were not exposed to damaging light levels at the surface to avoid any permanent structural and physiological damage to the photoreceptors even at low light levels (Figure 3.1). Between 7 and 25 individual fish of different sizes from above and below 500 metres (m) were collected from six localities (Table 3.1; see Figure 2.2). Individual specimens were removed from the light-tight net and transferred to a light-tight container and immediately taken to a darkroom and kept in iced seawater. Hereafter all procedures were performed in darkened rooms using only dim far-red illumination.

To examine the genetic variation underlying the visual system between the depth-associated *S. mentella* collections, we collected a total of 108 dorsal fin clips, to identify

SNP variation in the gene coding the visual pigment rhodopsin (as first reported by Shum *et al.*, (2014)), and 80 retinal tissues to examine the spectral sensitivity of *S. mentella*'s eyes at different depths. Additional genetic samples of *S. norvegicus* (n=3), *S. fasciatus* (n=31) and *S. viviparus* (n=4) were also analysed as outgroups. Eyes were hemisected on board and the retinae with attached retinal pigment epithelium were removed and frozen in seawater (adjusted to pH 6.5 using dd$H_2O$) at -20$^{o}$C.

**Figure 3.1** Illustrating the light-tight curtain covering the central codend of the multiSampler unit.

### 3.3.2 Generation of rhodopsin sequences and analysis

DNA was extracted using a modified salt extraction protocol (Miller *et al.*, 1998). An approximately 800bp fragment of the intron-free rhodopsin gene was amplified in 108 *S. mentella* samples using primers developed by Chen, Bonillo & Lecointre (2003): Rh193: 5'-CNT ATG AAT AYC CTC AGT ACT ACC-3' and Rh1039r: 5'-TGC TTG TTC ATG CAG ATG TAG A-3' and sequenced using Rh193. PCRs and temperature profiles followed Shum *et al.* (2014).

PCR products were purified and sequenced by Source BioScience Life Sciences (http://www.lifesciences.sourcebioscience.com/). Resulting sequences were manually checked for quality, edited and trimmed using CHROMAS LITE 2.1.1 (http://technelysium.com.au/?page_id=13) and aligned using MUSCLE v3.7 (Edgar 2004). In order to visualise the relationship between rhodopsin sequences, haplotype genealogies were constructed in the program HapView (Salzburger, Ewing & Haeseler 2011) using a maximum likelihood tree implemented in the program PHYLIP (Felsenstein, 1989).

### 3.3.3 Differential bleaching and analysis

In a dark laboratory, under dim far-red illumination, visual pigments were extracted and analysed using standard techniques (see Douglas *et al*., 1995 for details). In short, retinae were macerated in 1ml of either PIPES-buffered saline (PIPES 20 mM, NaCl 120 mM, KCl 19 mM, $MgCl_2$ 1 mM, $CaCl_2$ 1 mM; 300 mOsm/kg; pH 6.5). Then, 100 µl of 200 mM n-dodecyl β-D-maltoside, a mild detergent, was added to the homogenate and the solution agitated for 1hr at room temperature before being centrifuged for 10 mins (23000g, 4$^{o}$C).

The absorption spectrum (300-700nm) of 0.5ml of the dark adapted supernatant containing the visual pigment was determined using a Shimadzu UV2101-PC spectrophotometer following the addition of 30ul hydroxylamine ($NH_2OH$) (pH 6.5). This and all subsequent scans were zeroed at 700nm to ensure all pigment samples experienced equal light exposure prior to bleaching. The sample was then exposed to 3-5 mins of 501nm illumination at maximum light exposure, produced by a tungsten light source and a narrow band interference filters (10 nm bandwidth; B40 filters, Balzer, Liechtenstein), to bleach the visual pigment in order to determine the maximal absorbance ($\lambda_{max}$). A difference spectrum was constructed by subtracting the bleached scan from the unbleached spectrum and best fit with the visual pigment templates of Govardovskii *et al.* (2000), using methods described by Hart *et al*. (2000) to determine the wavelength of maximum absorbance ($\lambda_{max}$) of the visual pigment in the extract.

Five extracts, each from shallow (250-500m) and deep (550-800m) collections were subjected to more detailed bleaching at subsequent wavelength (see Figure 3.3a) and analysis (partial bleaching). Samples were exposed to a series of monochromatic bleaches of decreasing wavelength and rescanned between each bleaching event. If an extract contains two or more visual pigments, a difference spectrum of the initial dark scan and the first longwave bleach reveals an absorption spectrum dominated by the most longwave-sensitive visual pigment, whilst the difference spectrum of the last shortwave bleach reveals an absorption spectrum dominated by the most shortwave-sensitive visual pigment. If only one visual pigment is present in the extract, both difference spectra coincide (Knowles & Dartnall, 1977).

## 3.4 Results

### 3.4.1 Rhodopsin

We recovered four rhodopsin haplotypes defined by three variable sites across two shallow- and deep-pelagic clades (Figure 3.2). We observed a fixed non-synonymous amino acid substitution that separates the two clades, with the shallow clade exhibiting a **G**TC at position 119 (as referenced to the bovine rhodopsin: GenBank accession NM_001014890), which codes for Valine while the deep clade exhibits an **A**TC which codes for Isoleucine.

**Table 3.1** Location, Group code, and year of capture, trawl depth range, position, and number of individuals (N), Length: mean length in cm and length range for each collection of *S. mentella* investigated for 2013 and 2006/7 data from the North Atlantic. MS: multi-sampler.

| Location | Group code | Year | Gear | Trawl depth range | Latitude (N) | Longitude | N |
|---|---|---|---|---|---|---|---|
| **Irminger Sea** | | | | | | | |
| **Northeast deep** | $IrNE_{DP}$ | 2013 | MS | 550-650 | 62.44 | -29.08 | 25 |
| **Northeast shallow** | $IrNE_{SP}$ | 2013 | MS | 250-300 | 62.45 | -30.21 | 7 |
| **West shallow** | $IrW_{SP}$ | 2013 | MS | 200-275 | 61.44 | -38.35 | 18 |
| **West intermediate** | $IrW_{Int}$ | 2013 | MS | 475 | 61.45 | -40.19 | 16 |
| **West deep** | $IrW_{DP}$ | 2013 | MS | 600 | 61.44 | -38.53 | 20 |
| **Southwest shallow** | $IrSW_{SP}$ | 2013 | MS | 300 | 57.44 | -45.40 | 19 |

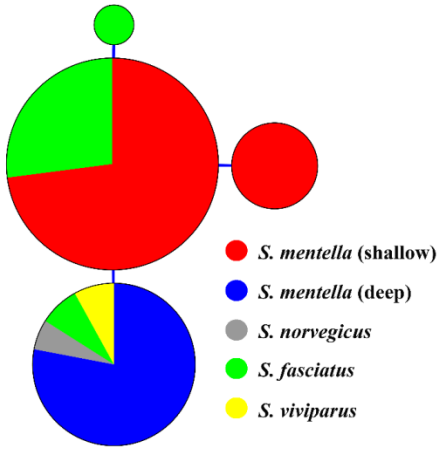



**Figure 3.2** Haplotype genealogy of rhodopsin sequences. The size of each circle represents the frequency of identifcal haplotypes and the lines interconnecting haplotypes represents one mutational step.

### 3.4.2 Visual spectra

Of the shallow and deep clades examined, all 5 partial bleaches revealed the presence of a single visual pigment and are accurately described well by the rhodopsin templates. (e.g. Figure 3.3c). In approximately 50% of the cases the fit between the extract difference spectrum and the visual pigment template did not meet quality control standards and the data had to be rejected. Table 3.2 shows the six collections studied, the number of retinae examined, the number that gave usable data and the average $\lambda_{max}$ of the fitted templates for all usable data in each group. A simple single factor ANOVA showed that $\lambda_{max}$ did not vary among different *S. mentella* collections, thus fish from all populations contain a visual pigment with indistinguishable absorbance properties, irrespective of depth and the amino-acid change at position 119. The average $\lambda_{max}$ of this pigment is 491.2nm.

**Table 3.2** Illustrating number of retinae examined. Sample codes refer to location and depth of each population (refer to Table 3.1 for location codes).

| | **Number of extracts** | | **$\lambda_{max}$** | |
|---|---|---|---|---|
| | total | usable | average | SD |
| **IrSW$_{SP}$13** | 21 | 12 | 491.8 | 0.7 |
| **IrNE$_{DP}$13** | 18 | 5 | 491.1 | 0.6 |
| **IrW$_{SP}$13** | 15 | 7 | 491.1 | 1.2 |
| **IrW$_{DP}$13** | 18 | 10 | 490.7 | 1.4 |
| **IrW$_{INT}$13** | 6 | 0 | - | - |
| **IrNE$_{SP}$13** | 2 | 0 | - | - |
| **All** | 80 | 34 | 491.2 | 1.1 |

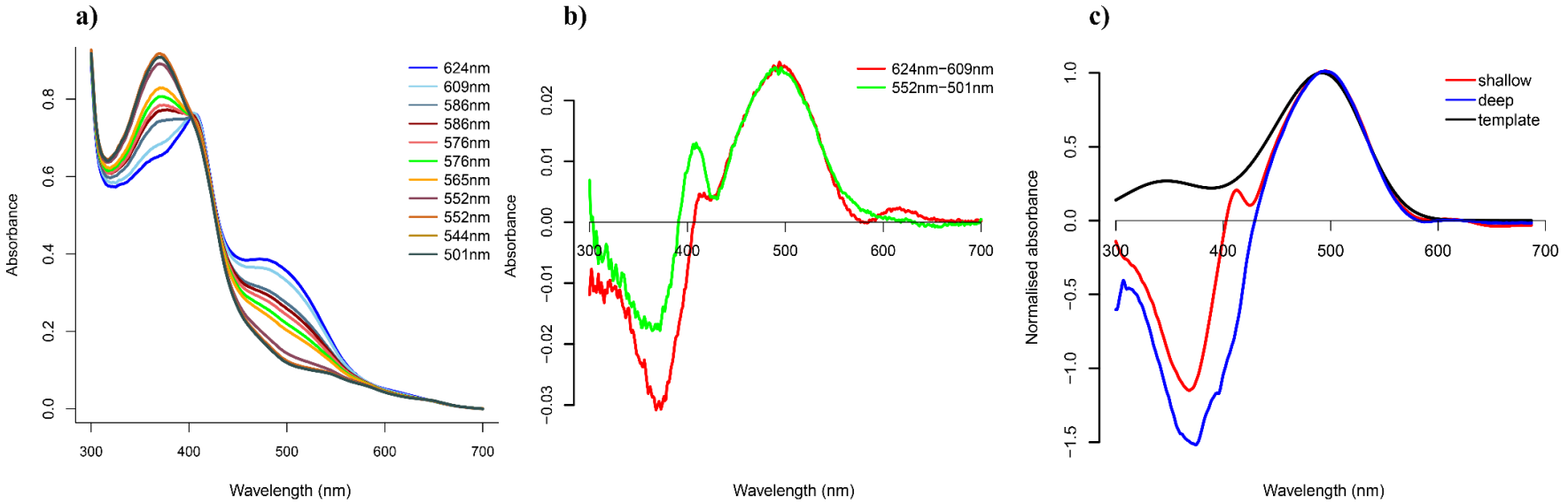



**Figure 3.3 (a)** Absorbance spectra of retinal pigment extracts from *S. mentella* following various bleaches of successively shorter wavelength. Initial measurement of the unexposed extract 10 min after the addition of 40 µl 1M hydroxylamine per 1 ml of extract; after eleven 5-min exposure intervals to monochromatic light of progressively shorter wavelength (see legend and colours). **(b)** Difference spectra constructed using the curves shown in (a): 624nm-609nm; 552nm-501nm. All have absorbance maxima around 491 nm (±<1nm), determined by rhodopsin template fitting, indicating the presence of only one pigment. **(c)** the total difference spectrum comparing the shallow and deep sensitivities against the best fitting template (491.9) of Govardovskii *et al.* (2000).

## 3.5 Discussion

Based on recent genetic evidence, we aimed to examine phenotypic consequences in depth-segregated groups of *Sebastes mentella* in the Irminger Sea. We present the first empirical attempt to link rhodopsin genotypes and visual pigment absorbance phenotypes in *Sebastes*, in relation to habitat depth. We find homogeneity in spectral sensitivity of rods across depth layers for the shallow- and deep-pelagic groups, corroborating Yokohama *et al.*'s (2008) expectation that rhodopsin amino acid changes at site 119 would not significantly affect spectral tuning. Although the present results fail to capture the paradigm of adaptive tuning of visual pigments by amino acid changes associated with depth, they lead to the generation of new hypotheses that may provide scope to study vision along a depth gradient.

### 3.5.1 Convergent evolution to depth

The genus *Sebastes* presents a system to study the evolution of traits that contribute to species divergence and colonisation to novel habitats (Alesandrini & Bernardi, 1999). Earlier findings based on molecular evidence showed that the gene encoding the visual pigment rhodopsin in North Pacific *Sebastes* appear to be adaptively evolving in relation to habitat depth (Sivasundar & Palumbi, 2010). This exemplifies the importance of studying the visual system of recently diverged species that colonised different photic environments, which can place constraints for prey acquisition, predator avoidance and mate choice (Smith *et al*., 2002; Smolka, Zeil & Hemmi, 2011; Meuthen *et al*., 2012). The analysis of *S. mentella* rhodopsin sequences expands on previous work with increased sampling (both individuals and localities) and supports the occurrence of a depth-defined SNP which strongly discriminates the shallow- and deep-pelagic mitochondrial clades (Figure 3.2; Shum *et al*., 2014). The SNP occurs at the first codon position and results in a fixed non-synonymous amino acid (AA) located within the transmembrane domain at site 119 (Figure 3.4). The shallow-pelagic clade exhibits a **G**TC which codes for Valine (Val119), whereas the deep-pelagic clade exhibits an **A**TC that codes for Isoleucine (Ile119). This pattern between the amino acid and depth habitat of *S. mentella* is mirrored in five North Pacific *Sebastes* species. Remarkably, four North Pacific species (*S. chlorostictus* (range 100-220m, depth change 31m), *S. elongatus* (100-230m, 2m), *S. aurora* (300-500m, 35m) and *S. melanostomus* (range 225-575m, depth change 35m))

associated with inferred depth changes to deeper waters exhibits an amino acid Ile119 as detected in the deep-pelagic clade. Similarly, one Pacific species (*S. diploproa* (range 200-325m, -10m)) has been linked to a secondary recolonization of shallower waters (Sivasundar & Palumbi, 2010) exhibits an amino acid Val119 and mirrors the polymorphism in the shallow-pelagic *S. mentella* clade. Since the Atlantic lineage diversified after the split from the Pacific groups (Hyde & Vetter, 2007), these patterns of genotype and depth preference are clearly independent events and therefore suggest a scenario of convergent evolution, a common feature of opsin evolution (Hunt *et al*., 1998). Light intensity is suggested to have a direct or indirect effect on depth distribution of *Sebastes* (Larson, 1980) and convergent evolution can explain the occurrence of repeat substitutions at certain sites associated with depth such as site 119. This would be reinforced if one could demonstrate that *S. fasciatus* (also bearing Val119) is adapted to shallower/clearer waters, however, considerably more work is required to elucidate similar patterns in this group.

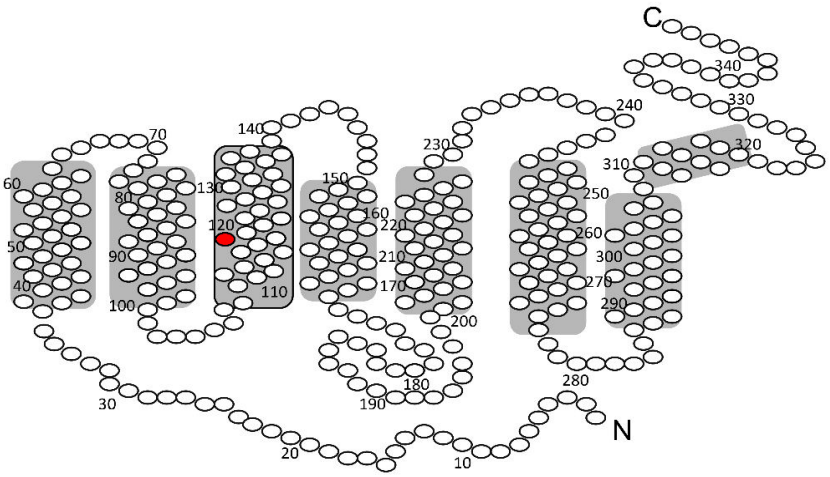



**Figure 3.4** Two-dimensional representation of the structure of rhodopsin. Position in red is site 119 located within the third transmembrane domain.

### 3.5.2 Spectral sensitivity

The amino acid changes occur in functionally important regions of the protein and were previously identified as being a target of positive selection which suggests a possible functional role for this amino acid site in *Sebastes* (Sivasundar & Palumbi, 2010; Shum *et al*., 2014). It is reasonable to hypothesise that the difference is a response to distinct light

spectra in depth environments. Yet, there seems to be no correspondence between the spectral absorbance properties of the shallow- and deep-pelagic *S. mentella* visual pigments as measured by spectrophotometry and the rhodopsin gene variation, concluding that the amino acid mutations at site 119 are not involved in spectral tuning. There are a number of inferences to be considered for this observation; 1) *S. mentella* are visual predators and hunt in overlapping photic environments and a thus exposed to similar conditions and 2) the non-synonymous mutations may be at the basis of evolutionary and molecular processes involved in receptor sensitivity such as phosphorylation.

Retinal extracts of shallow- and deep-pelagic specimens reveal a single rod visual pigment with the characteristics of vitamin $A_1$ based rhodopsin, with a peak spectral sensitivity of 491.2nm. It has been argued that the $\lambda_{max}$ of the rod pigments of deep sea fish (470-490 nm) are adapted to the spectral maximum of available light at depth where the spectral quality of light becomes progressively restricted to a narrow waveband of light in the ocean (470-480nm; Partridge *et al*., 1988; Douglas *et al*., 1997). However, although *S. mentella* dwells within the mesopelagic zone (200-1000m), their rod pigments absorb much light outside this narrow range of highest intensity, suggesting a potential adaptation to the detection of bioluminescent emissions at shorter or longer wavelengths ($\lambda_{max}$ = 450-510nm) produced by other deep-sea animals (e.g. Douglas *et al*., 1995; Turner *et al*., 2009). Gauthier and Rose (2002) used acoustic methods to study redfish vertical migrations in the Grand Banks of Newfoundland and report diel migration behaviour as a foraging strategy in pursuit of their euphausiid prey. Similarly in the scattering layers of the Irminger Sea, euphausiids (*Meganictyphanes norvegica*) were found to be one of the major components of the *S. mentella* diet (Magnússon and Magnússon, 1995; Pikanowski *et al*., 1999; Pétursdóttir *et al*, 2008). The relatively high abundance of *M. norvegica* in the *S. mentella* diet is noteworthy in light of a visual response because *M. norvegica* is one of the main emitters of bioluminescence in the scattering layers (Widder *et al*., 1992). However, *M. norvegica* are blue-green bioluminescent with a peak wavelength of 480nm (Craig & Preide, 2012) which is outside the peak sensitivity of *S. mentella*. Many fish species possess cones during early life stages as juveniles/larvae inhabiting surface waters and some are even retained in deeper living animals (Bowmaker *et al*., 1994). Therefore, perhaps, *S. mentella* possess the ability to exploit opsin genes that enable them to target their bioluminescent prey. However, an examination of the retina through histology and microspectrophotometry (MSP) would be required in order to directly verify the existence of such cells, and the tissue used for our analyses had deteriorated beyond suitable

conditions for analysis. Indirectly, transcriptomic analyses of expressed gene can also be employed to explore the existence of these functions, and this will be examined in the next chapter.

*S. mentella* are mobile predatory fish and their ability to undergo vertical migrations in the water column is perhaps a response to environmental conditions and prey availability. Positive vertical migrations (inhabiting deeper waters during the day and migrating towards shallower waters at night) has been previously reported for marine pelagic predators (tuna, Dagorn, Bach & Josse, 2000; sharks, Stevens *et al*., 2010; seabream, Afonso *et al*., 2014), and it is assumed this behaviour is an evolutionary driving force for developing morphological and physiological adaptations to provide increased visual abilities under low light levels (Dagorn, Bach & Josse, 2000). Afonso *et al*. (2014) document individual depth displacements upon monitoring vertical migration patterns for North Atlantic seabream, most likely a response to prey availability enacting this behaviour in low-light levels. Thus it may be expected that *S. mentella* groups display displaced/staggered vertical daily migrations (see Frank & Widder, 1997) and experience different light intensity. This behaviour may explain the comparable rod spectral sensitivities between the shallow- and deep-pelagic clades, while the staggered vertical distributions for fish exhibiting different light regimes support the rhodopsin genotypes generated over an evolutionary time scale, or simply the result of phylogenetic constraints (Douglas *et al*., 1997).

A closer examination of the phylogenetic history of North Atlantic *Sebastes* rhodopsin may serve to explain the differences between the two *S. mentella* clades. The haplotype genealogy illustrates the deep-pelagic group cluster with *S. viviparus*, *S. fasciatus* and *S. norvegicus*, suggesting the amino acid Ile119 is ancestral (Figure 3.2). Hence this pattern supports the pattern for a deep-pelagic clade stall/reduction in population growth during periods of low sea level and consequently retained the ancestral genotype (Shum *et al*., 2015). While the shallow-pelagic *S. mentella* has been shown to exhibit more pronounced population growth during the last glacial maximum, the emergence of an amino acid substitution may have followed allopatric conditions during periods of rising sea level, possibly resulting in a shift in receptor sensitivity (Shum *et al*., 2014). Ou *et al*. (2011) demonstrated that an amino acid substitution from Leucine to Cysteine (L119C) resulted in shorter meta-II lifetimes, which is an intermediate of rhodopsin involved in activating the visual G-protein, transducin. This substitution may influence the duration of receptor activity in the active state (i.e. meta-II) at the helix G-

protein binding site. Cysteine has similar properties to Isoleucine and Valine and therefore, if the mutations resulted in different rates of meta-II lifetimes, one could *hypothesize* that the receptor is more/less sensitive. However, it would not be possible to *conclude* that the receptor is more/less sensitive as there are also mechanisms (e.g. receptor phosphorylation and arrestin binding) that turn the receptor off and these processes may be happening at the same rates. Thus it can be argued that the amino acid variation may possess some adaptive value to light sensitivity in different optical environments to depth.

## 3.6 Conclusions

We provided a depth-associated examination of *Sebastes mentella* rhodopsin variation and, for the first time, described the visual pigments of *Sebastes* between the shallow- and deep-pelagic clades in order to link genotype with phenotype. The spectral absorption characteristics of the retinal photoreceptors revealed a single visual pigment, rhodopsin, with peak sensitivity at 491.2nm. We showed a fixed rhodopsin SNP, which results in an amino acid (AA) substitution associated with depth distribution of the shallow- and deep-pelagic types, which suggest that habitat use between these two Atlantic groups, might be attributed to convergent evolution with Pacific *Sebastes*. We find no mechanistic or evolutionary link between rhodopsin variation and the measured pigments. However, several inferences can be made about the visual ecology and rhodopsin molecular variation of *S. mentella*. *S. mentella* are likely to be adapted to hunting bioluminescent prey, possibly aided by cone opsin genes. Furthermore, although the SNPs are not involved in spectral tuning, there is evidence to suggest it may be involved in receptor sensitivity during phototransduction. Collectively, the present study strengthens the view of depth segregated population units in *S. mentella*, and also offers a springboard for further multidisciplinary analyses in this group of fishes.

# Chapter IV

# RNA-seq Reveals Adaptive Processes to Different Depth Environments in Beaked Redfish, *Sebastes mentella*

## 4.1 Abstract

Pelagic marine fish are exposed to a wide range of optical environments that affect the repertoire of visual genes they possess, from many deep-sea species exhibiting retinas with only rod opsin to shallower water species having multiple subfamilies of visual opsins. One candidate study system to examine adaptation to different visual environments is the vertically structured Atlantic beaked redfish, *Sebastes mentella*. We carried out *de novo* sequencing of the retinal transcriptome from a subset of "shallow" and "deep" *S. mentella* individuals, in order to characterise differential expression of candidate genes involved in sensory perception. We detected differential expression in 526 transcripts, 150 of which are upregulated and 376 are downregulated in the 'shallow' group. In particular, variation in gene expression involved in rhodopsin receptor sensitivity (phosphorylation), suggests a differential responsiveness for rhodopsin in fish adapted to different light environments. Unexpectedly for a species that dwells at significant depths, we provide the first transcriptomic evidence for colour vision driven by two families of cone opsins, suggestive of potential adaptation to bioluminescence or simply ancestral constraints from the diverse Sebastid radiation. These results offer further insights into the patterns and mechanisms of adaptation to the deep sea environment.

## 4.2 Introduction

The open ocean contains a vast diversity of organisms that have adapted to a wide range of optical environments and habitats (Lythgoe, 1988). Variation in depth, and consequently light levels, may result in adaptive divergence in phenotypic traits involved in vision (Douglas & Partridge, 1997). Thus, when assessing the potential

impact of differential photic regimes, it is important to determine whether species and/or populations separated by depth vary in their visual adaptations. Where such differences occur, insight into the underlying mechanisms is key to understand candidate genes responsible for population-specific adaptation to different light environments. Although numerous studies have examined the role of transcriptional phenotypic plasticity on the visual response for aquatic species in shallow waters (Hofmann *et al.*, 2010; Phillips *et al.*, 2015), little is currently known about the transcriptional mechanisms that influence visual perception in deep sea fish.

The live-bearing genus *Sebastes*, comprising more than 100 species, is the most highly speciose genus in the family Scorpaenidae (Hyde & Vetter, 2007). They can be found over a wide range of depths, from surface waters down to 1000m (Love *et al.*, 2002). Recent studies have found that many species have adaptively evolved in response to optical conditions at different depths (Sivasundar & Palumbi, 2010; Shum *et al.*, 2014) and it may be expected that *Sebastes* possess functional opsin genes involved in colour perception. The beaked redfish, *Sebastes mentella*, experience depth-related substructure in the North Atlantic (NA; Cadrin *et al.*, 2010). *S. mentella* occupy some of the most challenging optical conditions that occur within the mesopelagic zone (Ringelberg, 1995), with a shallow-pelagic group inhabiting depths between 200-550m throughout the NA while a deep-pelagic group dwell between 550-1000m within the Irminger Sea and Faroe Islands (Cadrin *et al.*, 2010; Shum *et al.*, 2015). Chapter III examined the pigment involved in dim-light vision, rhodopsin, and results show shallow- and deep-pelagic *S. mentella* possess a single rod visual pigment ($\lambda_{max}$=491.2nm) despite a fixed amino acid difference (site 119) that discriminates the two groups (see chapter III). However, the amino acid change was found to have negligible effect on the absorption spectra, which points to a physiological role in phototransduction (Ou *et al.*, 2013). Here, we focus on conspecific populations occupying a two depths environments of photic environments and assess differences in gene expression in identifying local adaptation to visual environments.

The visual response of marine fish depend on the absorption of available photons that allows formation of the best possible image in terms of contrast, movement and depth (Douglas & Djamgoz, 1990). Moreover, the presence of cone combinations adds further complexity to vision and enhances the perception of the environment. It is the visual pigments in the outer segments of retinal rod and cone

photoreceptor cells that absorb photons by 11-*cis*-retinal, which isomerizes to all-*trans*-retinal. This occurs through a phototransduction cascade, converting light into an electrical signal that leads to vision (Kashiyama *et al.*, 2009). The biochemical signalling of phototransduction involves a series of conformational shifts that lead to the active forms of metarhodopsin II (meta-II) which binds and activates transducin, the visual G-protein (Smith, 2010). In contrast, inactivation of a photoresponse involves a two-step process: first, G protein-coupled kinases (GRKs) phosphorylate active meta-II at multiple sites, thereby desensitising receptor/G protein activity; second, inactivation occurs when arrestin binds phosphorylated meta-II and terminates the remaining activity of the opsin (see Figure 1.3)

Changes in gene expression play an important role in visual sensitivity that drives adaptation in photic environments (Hofmann *et al.*, 2010). Marine fish show considerable variation in ontogenetic plasticity in both the combination and number of specific opsins expressed within the retina (Tekechi & Kawamura, 2005; Hofmann *et al.*, 2010; Valen *et al.*, 2013). Numerous studies document that visual pigments within the same cone classes may become activated or inhibited during development (Carleton *et al.*, 2005; Bowmaker *et al.*, 2008; Taylor, Loew & Grace, 2011; Flamarique, 2013). Thus, the visual sensitivity of an individual is influenced by external factors during its own life-history likely as a result of habitat shifts, prey preference, mate choice and available light acting on a subset of available opsins.

Here, we assess differential gene expression as a result of ecological and evolutionary adaptation of *S. mentella* to different depths in photic environments. In this study, we focus on genes related to visual perception due to their significance in controlling light detection in the deep sea. We collected retinal tissue from four shallow and four deep *S. mentella* individuals from different depths within the Irminger Sea, and then performed RNA-seq to gain a detailed snapshot of transcriptome-wide gene expression of visual related functions.

## 4.3 Materials & methods

### 4.3.1 Animals and sample collection

Four individuals for each of the two "shallow" and "deep" *Sebastes mentella* groups were collected at distinct depths in the Irminger Sea (Table 4.1; See Figure 2.2 for

sampling locations). Moribund animals were caught using a MultiSampler trawl fitted with a light-tight curtain on board the Icelandic r/s Arni Friðriksson from 6th of June – to 11th of July 2013. Individual specimens were immediately taken to a darkroom and put on iced sea water. The right eye of each animal was dissected and stored at -20˚C at sea and transported to the laboratory on dry ice, and then stored at -80˚C prior to RNA extraction.

**Table 4.1** Location, year of capture, trawl depth range, position, and number of individuals (N), for each collection of *S. mentella* for 2013 from the North Atlantic. MS: multi-sampler.

| Location | Year | Gear | Trawl depth range | Latitude (N) | Longitude | N |
|---|---|---|---|---|---|---|
| **Irminger Sea** | | | | | | |
| **Northeast deep** | 2013 | MS | 550-650 | 62.44 | -27.41 | 3 |
| **Northeast deep** | 2013 | MS | 550-650 | 62.44 | -29.08 | 1 |
| **Southwest shallow** | 2013 | MS | 300 | 57.44 | -45.40 | 4 |

### 4.3.2 RNA isolation, library preparation, and sequencing

Total RNA was extracted from frozen retinal tissue using the TRIZOL Purelink RNA mini kit, Ambion Life Technologies (Figure 4.1). RNA samples were then digested by DNase I to remove potential genomic DNA. The quantity and integrity of total RNA was assessed by Qubit fluorometer (Invitrogen, cat. no. Q32857) and the RNA Pico 6000 kit (Agilent 2100 Bioanalyzer), respectively. RNA-seq library preparation (N=8, 4 shallow and 4 deeps) and sequencing was carried out by the Liverpool Centre for Genomic Research (CGR) (https://www.liv.ac.uk/genomic-research/). Five micrograms of total RNA was ribosomally depleted using the Ribo-Zero Gold Kit (Epicentre) and purified with Agencourt RNAClean XP beads. Then, 50 ng of the depleted RNA was prepared into an indexed, strand-specific library using the ScriptSeq v2 RNA-Seq Library Preparation protocol. Following 10 cycles of amplification, eight libraries were purified using Agencourt AMPure XP beads. The final libraries were pooled in equimolar amounts, and sequenced on one lane using Illumina HiSeq 2500 at 2x100 bp paired-end reads.

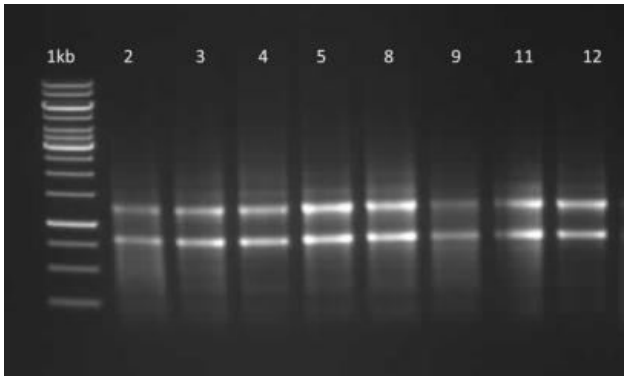



**Figure 4.1** Total RNA integrity of *S. mentella* retinal tissue on ethidium bromide-stained 1% agarose gel. Gel shows ribosomal RNA 28S (upper band, 1.2kb) and 18S (lower band, 0.8kb). mRNA species not observed but still exist below gel detectable levels and/or as a result to low exposure. Wells 2-5: shallow-pelagic, wells 8,9,11,12: deep-pelagic. 1kb: GeneRuler 1kb DNA ladder (Thermo Fisher Scientific).

### 4.3.3 Transcriptome assembly and functional annotation

The raw Fastq files resulting from the illumina HiSeq run were trimmed and Illumina adapter sequences removed using Cutadapt version 1.2.1 (Martin, 2011). The reads were further trimmed using Sickle version 1.200 (https://github.com/najoshi/sickle) with a minimum window quality score of 20. Reads shorter than 10 bp after trimming were removed. If only one of a read pair passed this filter, it was rejected.
*De novo* assembly of the *Sebastes* retinal transcriptome was carried out using the short read assembly program Trinity, using default parameters with 300 nt minimum contig length (http://trinityrnaseq.sourceforge.net/) (Grabherr *et al.*, 2011). The output of the illumina sequencing runs for all biological replicates were combined to maximise assembly accuracy. Reads of each biological replicate were mapped back to the assembled contigs using TopHat v2.0.5 software (Trapnell *et al.* 2009). Tophat is a mapping program specifically designed for RNA-seq reads that uses the fast Bowtie2 aligner and then determines splice junctions (Langmead & Salzberg, 2012).

Functional annotation was performed on the assembled contigs using the Trinotate annotation pipeline using default settings (http://trinotate.sourceforge.net/). First, we used Blastx alignment (Altschul *et al.*, 1990) against the protein SwissProt database. Second, open reading frames (ORFs) were predicted by the TransDecoder tool (Haas *et al.*, 2013) using default parameters, to identify candidate coding regions

within transcript sequences. Blastp (Altschul *et al.*, 1990) was used to search Transdecoder-predicted proteins against the SwissProt database and HMMER 3.1b2 (Finn, Clements & Eddy, 2011) to identify protein domains by searching the Pfam_A database (Punta *et al.*, 2012). The prediction of signal peptide and transmembrane domains was performed through SignalP (Petersen *et al.*, 2011) and tmHMM (Krogh *et al.*, 2001), respectively. Lastly, results were loaded into the Trinotate SQLite Database for comparison to EMBL Uniprot, eggNOG and the Gene Ontology (GO) Pathways annotation databases (Powell *et al.*, 2012; Kanehisa *et al.*, 2012). To characterize the protein composition of the *S. mentella* retinal transcriptome, a custom python script (available from http://github.com/NathanWhelan, last accessed October 10, 2015) was used to place GO terms for each UniProt annotated transcript from Trinotate into Web Gene Ontology Annotation Plotting (WEGO) format, and annotated GO terms were visualized using the WEGO web service (Ye *et al.*, 2006). Gene Ontology (GO) is used to standardize representation of genes across species and provides functional vocabularies for annotating genes, gene products, and sequences (Ashburner *et al.*, 2000).

### 4.3.4 Analysis of gene expression

To analyse differential expression patterns, quantification of raw read counts from alignment data were extracted using HTSeq-count from the HTSeq package (Anders, Pyl & Huber, 2014). Count data was used to find significant differences in transcript abundance between the shallow and deep collections of visual genes using the edgeR bioconductor package (Robinson, McCarthy & Smyth, 2010), which identifies differentially expressed transcripts based on the assumption that the number of reads produced by each transcript is used proportionally across sample replicates. EdgeR measures transcript abundance in counts per million (CPM) for each biological replicate and uses variance estimates to calculate the statistical significance (p-values) of observed differential expression. Count data were filtered prior to differential gene expression (DGE) analysis by removing transcripts with very low expression based on an expression cut-off of 3 CPM in at least four libraries. This is a filtering step to ensure that counts (genes) are removed that are not expressed. To account for multiple testing in differential expression analysis, p-values were adjusted to Benjamini and Hochberg's algorithm to control the false discovery rate

(FDR) (Benjamini & Hochberg, 1995). Selected genes from raw count values were normalised using housekeeping gene elongation factor 1 alfa (Manfrin *et al.*, 2015), and differential transcription was calculated as fold change expression between the shallow and deep group of visual genes.

## 4.4 Results

### 4.4.1 Transcriptome assembly and annotation

To obtain the *S. mentella* transcriptome expression profile, eight libraries were constructed using retinal tissue from each of the 4 shallow and 4 deep replicates. Each biological replicate produced a RNA Integrity Number (RIN) score of at least 7 or above which produced a total of 490,130,476 paired-end reads with an average paired-end read length of 2x100 bp among 8 replicates. After removing low quality or redundant reads and trimming for the presence of adaptor sequences, 479,898,444 clean reads were subsequently used for assembly and analysis. In total 479,898,444 reads were assembled into 450,735 isoforms with an N50 length of 1525 bp and a maximum length of 20,909 bp using the program Trinity. The length distribution of contigs is presented in Figure 4.2. In total, 210,028 contigs were ranging from 300 nucleotides (nt) to 2000 nt, with 55,427 contigs over 2000 nt.

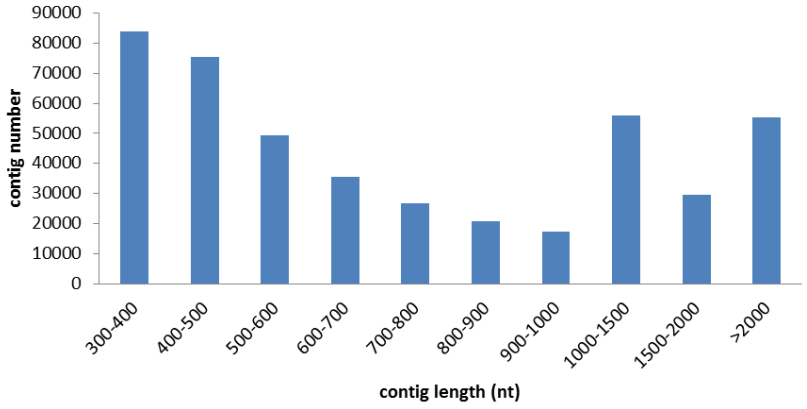



**Figure 4.2** Distribution of transcript contig length in RNA-seq data of *S. mentella*. nt: nucleotides

In total, 28% (n=127,984) of putative transcripts had predicted open reading frames (ORFs) and 80% (n=102,292) of inferred ORFs were annotated with a significant

BLASTp hit (identity hits, 1e-05). Of these annotated contigs, most had multiple Gene Ontology (GO) terms. The contigs were annotated in three gene ontology domains as biological process (32%, n= 99240), molecular function (31%, n=94637) and cellular component (36%, n=109891) (Ye *et al*., 2006). Within each of these three main GO categories, genes that annotated for cell part, organelle, binding, catalytic, cellular component organization, metabolic process and pigmentation were the most abundant (Figure 4.3).

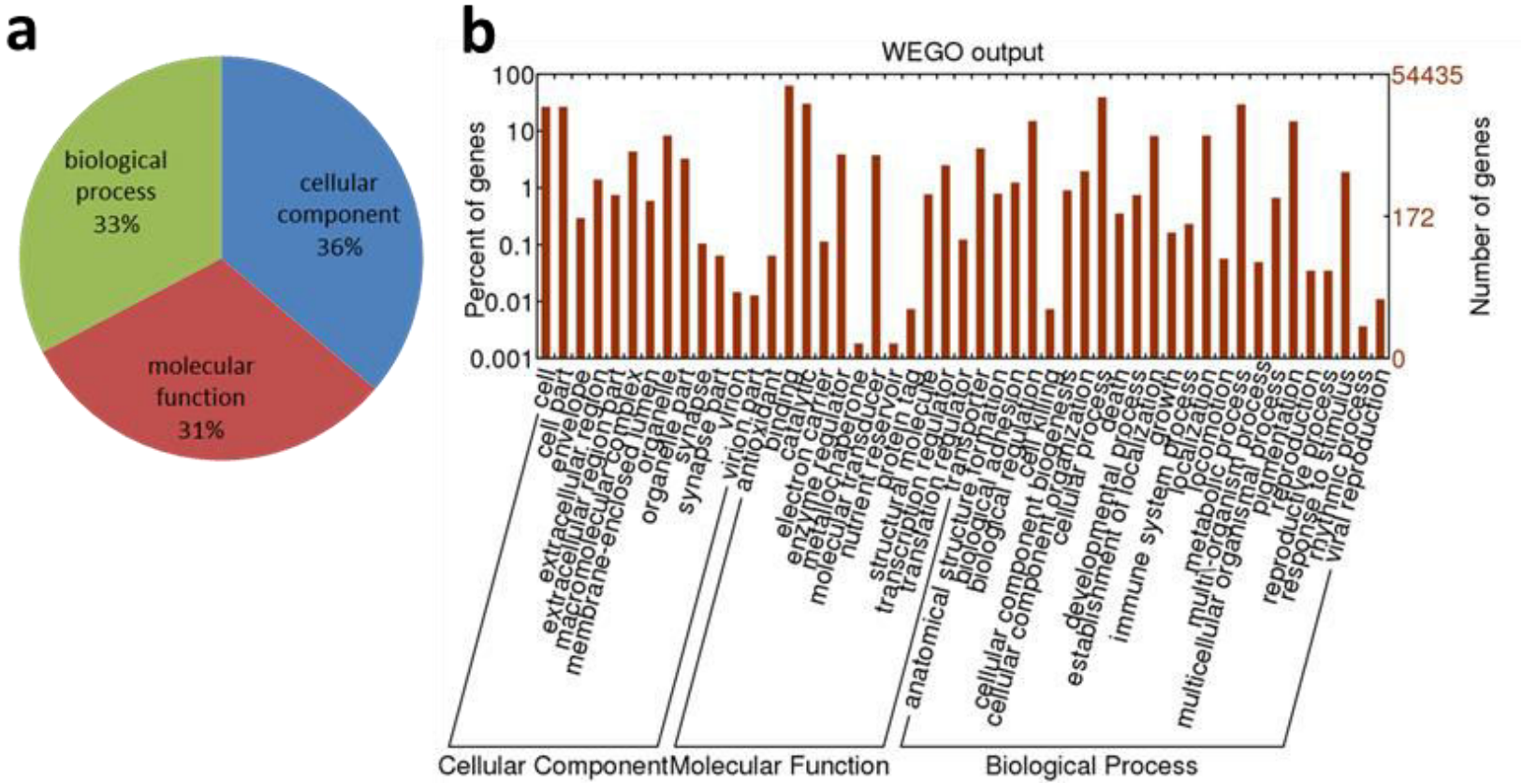



**Figure 4.3** (a) Pie chart with percentages of the higher-level GO terms for *S. mentella*, (b) Bar chart of genes annotated with select "Cellular Component", "Molecular Function" and "Biological Process" GO terms.

### 4.4.2 Differential expression

To identify candidate visual genes between the shallow and deep groups, we performed differential gene expression (DGE) analysis using edgeR on the retinal transcripts for the library size corrected read counts which had a minimum of three counts per million (CPM) across eight libraries from a total of 210,028 contigs. Between *S. mentella* groups, 526 (0.2%) genes were significantly differentially expressed (False discovery rate (FDR) <0.05), returning only 19% (n=103) of transcripts annotated. There were 150 transcripts significantly upregulated and 376 downregulated in the shallow group compared to the deep group (Figure 4.4).

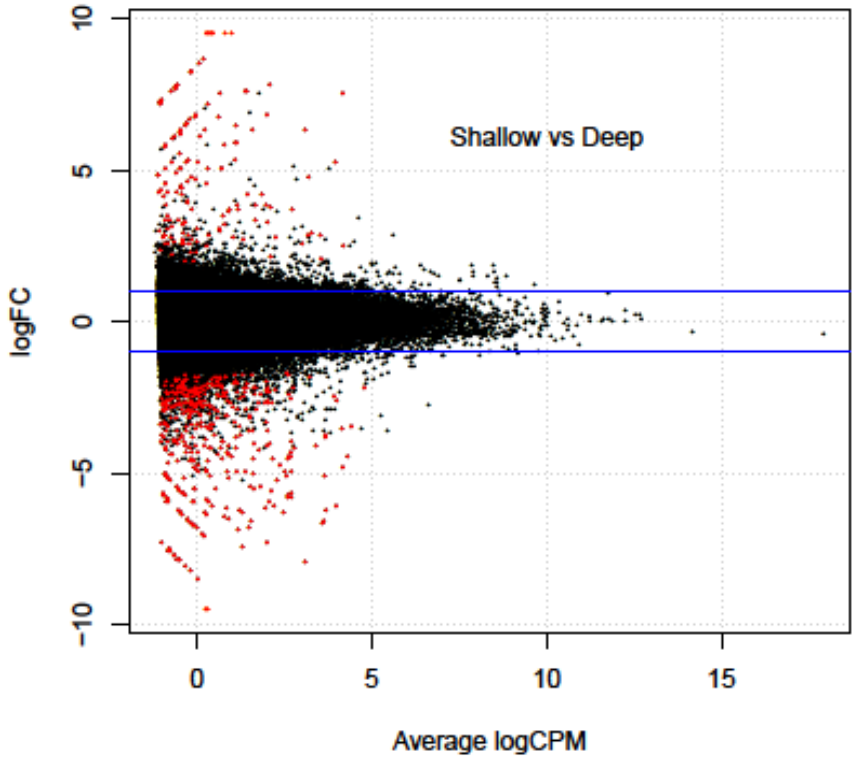



**Figure 4.4** Smear plot of expression of logarithmic fold change values (shallow vs. deep) against mean normalized read counts (Anders & Huber 2010). Red dots represent responding genes with a significant differential expression when using 5% false discovery rate (FDR) significance level.

### 4.4.3 Visual genes

Putative genes involved in visual perception were recovered from the *S. mentella* transcriptomes (Table 4.2; Figure 4.5). We identified five opsins in 13 assembled

transcripts (i.e. isoforms). These genes were among co-regulated and differentially expressed genes involved in dim-light vision (*rhodopsin*), colour perception (*blue-sensitive opsin*, *green-sensitive opsin*, *ultraviolet-sensitive opsin* and a *putative violet-sensitive opsin*), and genes that participate in light-induced cascades (e.g. G protein coupled-receptor kinase 4 (GRK4) and arrestin-C) that influence behaviour and physiological processes (Ou *et al.*, 2013). Among the four types of cone opsins detected in *S. mentella*, the green-sensitive opsin showed the highest expression on average followed by the blue-sensitive opsin, while both the ultraviolet-sensitive opsin and putative violet-sensitive opsin show a lower levels of expression compared to housekeeping gene EF-1α (Figure 4.5). Blue-sensitive opsin and arrestin-c was found to be significantly upregulated in the shallow group. GRK4 was found to be significantly downregulated in the shallow group which phosphorylates the activated forms of rod G protein-coupled receptors. Among the co-regulated genes, rhodopsin and green-sensitive opsin are highly expressed relative to housekeeping gene EL-1α. Furthermore, we identified candidate genes involved in adaptive immunity including major histocompatibility antigen (MHC class I & II), Interleukin-8 (IL-8), B-cell lymphoma 3 (Bcl-3) and immunoglobulin heavy chain (Ig heavy chain) all found to be downregulated in the shallow group (Table 4.2).

**Table 4.2** Repertoire of genes associated with vision and immunity. Average fold change relative to housekeeping gene Elongation factor 1 (alpha EF1-α).

| Family | Gene name | Abbr | Gene ID | Fold increase | |
|---|---|---|---|---|---|
| | | | | shallow | deep |
| **Co-regulated** | | | | | |
| **UVOP** | Ultraviolet-sensitive Opsin | UV | TR171708\|c0_g1_i1 | 0.14 | 0.09 |
| **OPSV** | Putative violet-sensitive Opsin | V | TR 79966\|c0_g1_i1 | 0.25 | 0.05 |
| **OPSG** | Green-sensitive Opsin | GO | TR 186358\|c0_g1_i1 | 3970 | 1858 |
| **Up-regulated** | | | | | |
| **OPSB** | Blue-sensitive Opsin | BO | TR 29672\|c0_g1_i1 | 9 | 1.3 |
| **CAR** | Arrestin-C | CAR | TR 39788\|c0_g1_i2 | 60 | 12.5 |
| **Down-regulated** | | | | | |
| **GRK4** | G protein-coupled receptor kinase 4 | GRK4 | TR 32159\|c7_g1_i1 | 2.8 | 11.7 |
| **HLA-C** | MHC class I antigen Cw*8 | HLA-C | TR 449\|c0_g4_i1 | 57.7 | 375.8 |
| **H2-L** | MHC class I antigen, L-D alpha chain | H2-L | TR 40919\|c8_g2_i4 | 36.7 | 92.2 |
| **B-F-beta-IV** | MHC Class I antigen, F10 alpha chain | B-F | TR 60388\|c1_g1_i1 | 8.3 | 220 |
| **Mamu-DRA** | MHC class II antigen, DR alpha chain | DRA | TR 41037\|c2_g1_i2 | 13.5 | 139.8 |
| **Bcl-3** | B-cell lymphoma 3 | Bcl-3 | TR 14677\|c12_g2_i2 | 11.2 | 39.8 |
| **CXCL8** | Interleukin-8 | CXCL8 | TR20178\|c0_g1_i1 | 3.6 | 14.9 |
| **Ig** | Immunoglobulin heavy chain V-III | Ig | TR35775\|c0_g1_i1 | 0.75 | 2.3 |

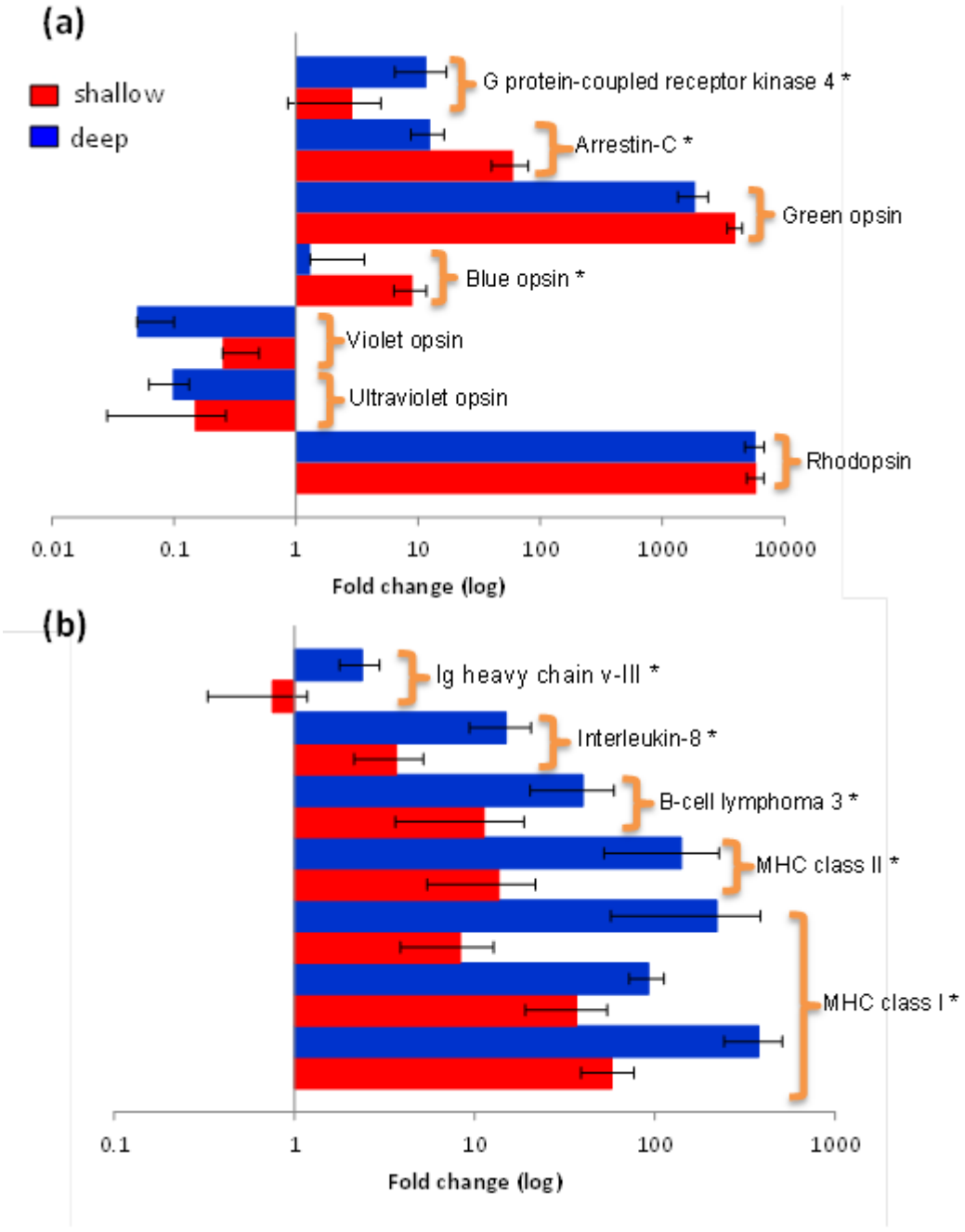



**Figure 4.5** Expression fold changes of selected retinal transcripts. Selected genes involved in (a) in visual processes, (b) immunity. Y-axis shows the expression level normalized on that of the elongation factor EF1-alpha. * indicates genes differentially expressed in DGE analysis. See Table 4.2 for fold-change values.

## 4.5 Discussion

Here we report the retinal transcriptome of *Sebastes* spp. using RNA-seq between the "shallow" and "deep" *S. mentella* in the Irminger Sea. *De novo* assembly generated 450,735 transcripts and identified the differential expression of 526 genes between shallow and deep groups. *S. mentella* exhibits gene expression patterns in rods (rhodopsin), cones (blue and green) and along the visual pathways that can be used to identify related visual functions. We identified a receptor kinase linked to differential responsiveness for rhodopsin between the *S. mentella* groups, and unexpectedly for species that dwell at significant depths, discovered functional cone opsin genes (Figure 4.5a) involved in photopic vision that allows colour perception. Differences in functional genes involved in vision hint that their expression may depend on life history strategies in response to different abiotic and biotic factors and/or evolutionary history of each group.

### 4.5.1 Rhodopsin sensitivity

*S. mentella* displays a single rod visual pigment with an amino acid difference in the third helical domain at site 119 that discriminates the shallow and deep groups (Shum *et al.*, 2014; Chapter III). Although the amino acid change at this position is not involved in spectral tuning, indicated by similar spectral sensitivities in the two groups, Ou *et al.*, (2013) showed an amino acid substitution L119C resulted in shorter meta II lifetimes. This suggests shorter meta-II lifetimes maybe a correlate that affect sensitivity of a photoresponse with membrane proteins such as the visual G protein, G protein-coupled receptor kinases (GRKs) and/or arrestin (see Arshavsky, 2002 for review). Nevertheless, rhodopsin deactivation through phosphorylation (e.g. mediated by GRKs) and/or arrestin binding is the only mechanism relevant to setting photoresponse duration in rods (Arshavsky, 2002). In general, the more receptor kinase present, the shorter the receptor lifetime in the active state and, therefore, the sensitivity of any given response is reduced.

GRK1 have been identified as the key participant to specifically phosphorylate active G protein-coupled receptors (meta II), leading to desensitisation of receptor activity (Premont *et al.*, 1995). However, GRK4, one of six GRK subtypes, has four splice variants (GRK4α, −β, −γ, −δ) of which GRK4α is shown to

bind and phosphorylate activated meta II in vitro (Sallese *et al*., 1997), although the functional role in vivo remains uncertain. GRK4 is found to be downregulated in the shallow compared to the deep group. Differential expression of GRK4 may impact phosphorylation rates triggered by a physiological adaptive response in the rod sensory receptors. The rod receptors would likely respond to GRK4 at different levels than would be possible if absorption sensitivity could not adjust in response to varying stimuli. The occurrence of GRK4 appears reduced in the shallow group, which may imply the receptor is less sensitive to desensitisation as they are exposed to greater ambient light levels compared to the deep group. This physiological mechanism may have evolved during the Pleistocene glaciations, underpinning adaptation of upper water layers in the "shallow" *S. mentella* clade (Shum *et al*., 2015).

### 4.5.2 Opsin expression

Marine fish are exposed to a wide range of optical surroundings and the number of cone classes present in each species may vary accordingly to both ontogeny and the photic environment (Valen *et al*., 2013). It has been predicted that *Sebastes* possess cone opsins in which some cones may have been lost functionality given their decreasing importance to darker environments (Sivasundar & Palumbi, 2010). We identified four types of cone opsins (*blue-sensitive opsin*, *green-sensitive opsin*, *ultraviolet-sensitive opsin* (UV) and a *putative violet-sensitive opsin* (V)) through *de novo* transcriptome assembly. Transcriptome analysis reveal that *S. mentella* have low levels of expression ("non-functional") for genes coding for opsins sensitive in the UV/V field of the light spectra, while the blue- and green-sensitive opsins exhibit functional levels at different magnitude. Light rapidly changes along the water column in which UV/V light is present in the upper epipelagic zone (0-100m), whereas the blue-green light can extend into mesopelagic environment. Green light is extinguished in the upper mesopelagic zone while blue light extend to greater depths and become progressively restricted to a narrow waveband of light (470-480nm; Douglas & Partridge, 1997). Thus the loss of UV/V sensitive opsin function and gene expression patterns in blue- green-sensitive cones, including rhodopsin, may indirectly elucidate a behavioural link between present opsins and the habitats *S. mentella* commonly occupy. For example, Flamarique (2013), using rainbow trout,

showed that UV cones are important for foraging juvenile fish, however as the fish grows and changes prey type, the blue cones become more favourable.

Ontogenetic plasticity in expressed cone opsins is a common feature in several species and the loss of UV cones in adult teleosts has been reported in a number of species (Hawryshyn *et al.*, 1989; Kunz *et al.*, 1994). Adult *S. mentella* have dichromatic pigment genes (blue and green) and rhodopsin, whose expression patterns differ greatly. The green opsin and rhodopsin display, on average, a 490 and 990 higher fold changes compared to the blue opsin respectively. The dominance of green- to blue-sensitive cones has been documented in several studies (Valen *et al.*, 2013) and is shown to spread evenly throughout the retina (Helvik *et al.*, 2001). Thus, this can explain the greater expression levels observed in *S. mentella* green-sensitive opsin expression.

Furthermore, variable light conditions become distorted with depth and can consequently affect opsin sensitivity and gene expression (Dalton *et al.*, 2015). Blue-sensitive opsin is differentially expressed and upregulated in shallow *S. mentella*. When blue-sensitive opsin expression level is compared to that of the housekeeping Elongation factor 1 alpha transcript (EF 1-α), the shallow group show a significantly higher expression levels (9-fold), while no significant fold change is observed for the deep group. This indicates a higher sensitivity in blue opsin genes and/or increased exposure to the blue spectrum ($\lambda_{max}$ = 360-430nm; Yokoyama 2000) for shallow individuals dwelling in the upper layers of the mesopelagic zone.

Cone arrestin (Arrestin-C), involved in terminating light activated red/green cone opsins (Zhu *et al.*, 2003), was significantly upregulated (4.8-fold) in the shallow group. Renninger *et al.* (2011) found that functional loss of cone arrestin causes a prolonged photoresponse recovery and reduced temporal contrast sensitivity in zebrafish larvae. This implies *S. mentella* possess differential contrast sensitivities in green-sensitive opsin which largely depend on the exposure to longer wavelengths. The green spectrum of light is limited in the upper water layers and perhaps offers increased exposure to shallow *S. mentella*. This suggests the photoresponse recovery and contrast sensitivity is greater for fish occupying the upper water column, hence depth likely poses a strong impact on redfish retina.

The ability of possessing functional cone opsins remains uncertain for *S. mentella* inhabiting dim-light environments. It is unknown whether the regional matching of sensitivity to the visual field is developmentally plastic or under strict

genetic control (Carleton *et al*., 2010; Hofmann *et al*., 2010). Redfish exhibit diel migrations, which increase or limit their exposure of absorbing light at shorter or longer wavelengths (Gauthier & Rose 2002). Moreover, they are known to consume bioluminescent prey that emits a blue-green colour (Magnusson and Magnusson, 1995; Pétursdóttir *et al*, 2008; Widder, 2010), thus possessing a visual advantage increases the probability of capturing prey in low light environments. However, *S. mentella* opsins are likely under genetic control and maintained due to phylogenetic constraints. The genus *Sebastes* is composed of species that are morphologically and ecologically divergent and are notable for their diversity in form and function. The centre of this diversity located in the Northeast Pacific (Hyde & Vetter, 2007), and *Sebastes* spp. experience a wide range of light environments and depths from near surface waters to an excess of 1000m (Love *et al*., 2002). Thus, colour perception likely has an important impact on *Sebastes* behaviour that allows detection of optical signals from conspecifics (e.g. mate recognition, mate choice), suggesting a role for both natural and sexual selection in colour vision (Sabbah *et al*., 2010). However, the expression of an opsin may depend greatly on a limited area of the retina for a specific visual function (Takechi & Kawamura, 2005).

### 4.5.3 Other important findings: Immune response

We identified a repertoire of genes involved in innate and adaptive immune response including B-cell lymphoma 3 (Bcl-3), immunoglobulin heavy chain (Ig heavy chain) and major histocompatibility antigen (MHC class I & II), Interleukin-8 (IL-8), that were downregulated in the shallow group. This is of principal interest because parasites and abnormalities in pigmentation are routinely used as biological tags to discriminate the *S. mentella* groups, in which the shallow group present a higher rate of infection (Magnusson & Magnusson, 1995; Marcogliese *et al*., 2003). Bcl-3 is a subtype of the IκB family and is evolutionarily conserved and coordinates gene transcription in organisms in response to infection, stress and injury as a first barrier of defence against pathogens (Baeuerle & Henkel, 1994; Ghosh *et al*., 1998). Furthermore, Ig heavy chain is involved in the production of specific antibodies to a variety of antigens found in teleosts (Alcorn & Pascho, 2002; Solem & Stenvik, 2006). MHC I & II proteins, involved in presenting foreign antigens to the immune system, was accompanied by upregulation of IL-8, a chemokine essential for

recruitment and activation of T-cells through chemotaxis (Gesser *et al*., 1996). MHC are involved in many important biological traits and their variation and diversity make MHC a candidate gene to study the mechanisms and significance of molecular adaptation in vertebrates (see Sommer, 2005 for review). Wegner *et al*. (2003) examined an association between MHC variation and the number and diversity of parasites affecting different stickleback populations and concluded MHC diversity was positively correlated with parasitic load. The variance in parasitic infection is routinely used as a biological tag to discriminate shallow and deep *S. mentella* and retinal transcription highlight important aspects of immune-related proteins and their biological role in immunity. Deep *S. mentella* appears show higher expression patterns to infection and pathogens which simply reflect the harsh conditions associated with depth (Somero, 1992; Ringelberg, 1995) and so avoid visual impairment given the light-induced stress in darker environments.

## 4.6 Conclusions

Here we carried out the first *de novo* retinal transcriptome assembly of *S. mentella* to date and describe important genes that play an important biological role. The annotated genes provide an important toolkit for future avenues on vision and immune response in *S. mentella* as well as other *Sebastes* species. We identified 526 differentially expressed transcripts, of which 150 transcripts were significantly upregulated and 376 downregulated in the shallow group compared to the deep group. In particular, the data from this study reveal *S. mentella* did not express opsin genes in the UV/V light spectra while maintaining functional cone opsin in the blue and green spectrum. This can be attributed to the deep water habitats and/or bioluminescent prey they consume. Furthermore, physiological processes such as phosphorylation and arrestin binding, involved in mediating vision are associated with increased sensitivity in the shallow group. GRK4, previously implicated in rhodopsin phosphorylation in vitro, indicate the shallow group are subjected to a longer photoresponse, possibly as a result of greater ambient light levels. In addition, expression patterns of blue-sensitive opsin and cone arrestin maybe the result of increased exposure to shorter and longer wavelengths in the blue and green light spectra respectively. Moreoever, we identified key immune-related proteins which may impede deep *S. mentella* retinas from being compromised against pathogens,

stress and/or injury. The findings in this study provide a platform to stimulate future comparative genomic investigations of *Sebastes* species adapted to different ecological conditions that affect vision and immune responses.

# Chapter V

# MtDNA Approach Resolves Ambiguity of "Redfish" Sold in Europe

## 5.1 Abstract

Morphologically-based identification of North Atlantic *Sebastes* has been controversial and misidentification may likely produce misleading data, with cascading consequences that negatively affect conservation and seafood labelling. North Atlantic *Sebastes*, comprises of four species, commonly known as "redfish", but little is known about the number, identity and labelling accuracy of redfish species sold across Europe. We used a molecular approach to identify redfish species from "blind" specimens to evaluate the performance of the Barcode of Life (BOLD) and Genbank databases, as well as carrying out a market product accuracy survey from retailers across Europe. The conventional BOLD approach proved ambiguous, and phylogenetic analysis based on mtDNA control region sequences provided a higher resolution for species identification. We found the presence of two species of redfish (*S. norvegicus* and *S. mentella*) and one unidentified Pacific rockfish marketed in Europe. However, public databases revealed the existence of inaccurate reference sequences, likely stemming from species misidentification from previous studies, which currently hinder the efficacy of DNA methods for the identification of *Sebastes* market samples.

## 5.2 Introduction

Assessing the state of marine resources often requires international effort to monitor how fish populations change over time and examine historical trends to determine distribution and status of fish stocks. This involves gathering catch and fishery-independent survey data, interdisciplinary evaluation on the biology and connectivity of populations (Cadrin *et al*., 2010; Planque *et al*., 2013), as well as monitoring of market products (Miller *et al* 2012; Watson *et al*., 2015). The growing global importance of seafood trade, and the

parallel advances in food technology, processing and packaging techniques, and the complex supply networks make it necessary to ensure the authenticity and origin of seafood products (Marko *et al.*, 2004). Despite European Law (EC No. 2065/2001), which requests appropriate seafood traceability and labelling, the identification of species is often problematic because morphologically similar species are difficult to separate by anatomical characters. As a result, resource management can be negatively affected through inflation of catch data for more desirable species and underreporting for less desirable species.

North Atlantic *Sebastes* is comprised of four closely related species commonly referred to as "redfish": the beaked redfish *Sebastes mentella* Travin 1951, the golden redfish *Sebastes norvegicus* (previously known as *S. marinus*) Linnaeus 1758, the Acadian redfish *Sebastes fasciatus* Storer 1854 and the Norwegian redfish *Sebastes viviparus* Krøyer 1845. They are ovoviparous (i.e. internal fertilisation), long-lived, slow-growing, late-maturing, and have in general low natural mortality which makes them vulnerable to even low levels of harvesting (Planque *et al.*, 2013). All species are commercially exploited and have experienced decreasing fishery landings since mid-1990 (Marine Research Institute, 2014; Pauly and Zeller, 2015). The identification of these species is controversial and remains difficult due to overlapping meristic and morphological features, which leads to these species being often marketed under a single vernacular name, "redfish". The practice of marketing species under an 'umbrella' name has important consequences because species with different conservation needs are mixed together, which, in turn, compromises the ability of consumers to make informed responsible purchasing decisions (Griffiths *et al.*, 2013).

Seafood authentication and traceability has led to a recent surge in molecular-based approaches for species identification. DNA barcoding is a rapidly growing technique for species identification in which DNA sequence profiles of the mitochondrial cytochrome c oxidase I (COI) for a vast range of animal species is deposited in the Barcode of Life (BOLD) database, and can subsequently be used to correctly identify unknown specimens (Hebert *et al.*, 2003). There are several documented cases that illustrate the power of this approach as a tool to detect seafood mislabelling, whereby fish products can be potentially disguised for less desirable, cheaper or more readily available species (Miller & Mariani, 2010; Miller *et al.*, 2012). However, the BOLD data system has a limited amount of reference vouchers for some species and presents a challenging case for species that have recently diverged (Steinke *et al.*, 2009). When such cases require clarification, rapidly

evolving molecular markers offer additional DNA-based tools to distinguish between species (Vinas & Tudela 2009). Mitochondrial genes are limited to only the matrilineal lineage, which can subsequently hinder the correct identification of a species, particularly when hybridisation is common (Pampoulie & Daníelsdóttir, 2008). However, Shum *et al.*, (2015) have shown the mtDNA control region to show high resolution to distinguish *S. mentella* group with only 4% of mismatches between mtDNA and nDNA. GenBank is a comprehensive database that contains publicly available sequences for more than 300,000 organisms, across any portion of their genomes. It locates regions of similarity between sequences, producing a list of matches most similar to the query sequence and an estimate of the percentage identity. Although less rigorously monitored, the GenBank data base therefore provides an additional platform to match unknown sequences to their most closely related taxa.

Here we applied a mtDNA approach to assess the seafood trade in North Atlantic *Sebastes*, by conducting an investigation into the patterns of redfish marketed from various retailers across Europe. Specifically, we sought to 1) compare the BOLD and Genbank databases in order to evaluate their performance in identifying blind and market specimens for the mtDNA COI and control region (d-loop) respectively, 2) determine the identity of North Atlantic *Sebastes* species being marketed in a selection of European retailers. Results reveal that for some of these recently diversified species the classical BOLD approach remains ambiguous, and that bias in morphological identification from previous *Sebastes* studies has been deposited into the GenBank data base, leading to the incorporation of incorrect reference sequences.

## 5.3 Materials and methods

### 5.3.1 Sample collection

We targeted North Atlantic *Sebastes* "redfish" species from randomly chosen fish mongers or fish counters from supermarkets across Europe (Table 5.1). Fin clips were taken from whole bought specimens or muscle tissue from fillets and initially stored in silica before being immersed directly in 100% ethanol and stored at -20˚C. In order to further validate our approach to the identification of North Atlantic *Sebastes* species, we also analysed 29 blind specimens of known origin that were inspected visually (hereinafter referred to as

"MarRef"), shared by the Marine Research Institute in Iceland, and used the multimarker data set in Shum *et al.* (2015) as a reference custom data base.

### 5.3.2 Molecular analysis

DNA was extracted using the DNeasy kit (Qiagen ©) following the manufacturer's protocol. Each sample was amplified at two markers, the universal 650bp DNA barcode cytochrome c oxidase subunit I (COI; Ward *et al.*, 2005) and the mitochondrial control region (d-loop) 573bp fragment, using Hyde and Vetter's (2007) *Sebastes*-specific primer pair. PCRs and temperature profiles for d-loop genes follow Shum *et al.* (2014) and COI follow Serra-Pereira *et al.*, (2010). PCR products were sequenced using the Thr-Rf for d-loop and R1 for COI by SourceBioscience and sequences edited and trimmed using CHROMAS LITE 2.1.1 (http://technelysium.com.au/?page_id=13) and aligned with MUSCLE implemented using MEGA v.6 (Tamura *et al.*, 2013).

### 5.3.3 Data analysis

The COI sequences were identified using the Barcode of Life Data Systems online (BOLD, Biodiversity Institute of Ontario, University of Guelph, Guelph, Ontario, Canada; www.barcodinglife.org; Ratnasingham and Hebert 2007) using the species-level barcode database to identify each sequence. The d-loop sequences were cross-referenced using BLAST on GenBank (Basic Local Alignment Search Tool, National Centre for Biotechnology Information, Bethesda, Maryland; www.ncbi.nlm.nih.gov/). A threshold of 99% to 100% sequence similarities was used above which identification of unknown samples was deemed reliable.

#### *5.3.3.1 Sequence analysis*

Market and blind samples were analysed along with a subset of previously characterised samples of North Atlantic *Sebastes*, which included *S. viviparus, S. fasciatus, S. norvegicus* as well as shallow and deep *S. mentella* types (Shum *et al.*, 2015). Sequence identification was strengthened via phylogenetic analysis, nesting market and blind samples together with a subset of reference samples from Shum *et al.*, (2015). Furthermore, all publicly available published sequence data of North Atlantic *Sebastes* was

included from Genbank as a further source of reference that derived from two more studies. First, we obtained COI and d-loop data generated by Hyde and Vetter (2007) for all four NA *Sebastes*, including selected Pacific species (see results), which was used to investigate the evolutionary relationships of over 100 *Sebastes* species. Second, we obtained d-loop data generated from Artamonova *et al*. (2013) that was used to investigate hybridisation and diversification of redfish in the Irminger Sea. Phylogenetic relatedness among sequences was reconstructed in MEGA version 6 (Tamura *et al* 2013) using the neighbor-joining algorithm (Saitou & Nei, 1987), and the Kimura 2-parameter distance. The NJ was used as it is a simple approach to determine the position of *Sebastes* "products" within "species clusters". All positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons (Pairwise deletion option). Evaluation of statistical confidence in nodes was based on 1,000 non-parametric bootstrap replicates (Felsenstein, 1985).

## 5.4 Results

### 5.4.1 Marker diversity

A total of 48 samples were analysed of which 19 were collected from 12 different retailers from four European countries (Belgium, Germany, Norway and the UK) and the remaining 29 blind samples were obtained from the Marine Research Institute in Iceland. The comparison of interspecific genetic variability for all market and blind samples considered indicated the mtDNA d-loop (0.018 ± 0.003) has about a 3-fold greater nucleotide diversity than the COI (0.006 ± 0.001) marker.

### 5.4.2 Sequence identification

Sequences were initially screened using the DNA sequence databases BOLD and BLAST to identify the closest matching sequences for COI (420-651bp, average length 544bp) and d-loop (381-506bp, average length 486bp) respectively. The BOLD search (99%) for market and blind samples resolved only 15% (N=6) of the matches allowing for confident assignment of species, while 85% (N=33) returned ambiguous results due to high sequence similarity with more than one redfish species or could not be reliably identified at species level (e.g. *Sebastes* spp.). The BLAST search of *Sebastes* d-loop sequences returned 56% (N=21) of the matches identified at species level while 44% (N=27) of the matches produced ambiguous results (Table 5.1, 5.2).

**Table 5.1** Market samples analysed for species identification with label information and results from BOLD and Genbank databases using the mtDNA COI and control region respectively. Species i.d. is inferred from phylogenetic reconstruction based on d-loop sequences (see Figure 5.2).

| Code | sold as | Country | BOLD database (COI) >99% | Genbank (D-loop) >99% | Species i.d. (d-loop) |
|---|---|---|---|---|---|
| **Bergen_1** | redfish | Norway | *S. marinus* (also *S. viviparus*) | *S. fasciatus* / *S. marinus* / *S. mentella* | *S. norvegicus* |
| **Bergen_2** | redfish | Norway | *Sebastes* sp. / *S. mentella* / *S.marinus* | *S. marinus* / *S. mentella* | *S. mentella (deep)* |
| **Oslo_3** | redfish | Norway | *Sebastes* sp. (also *S. marinus*) | *S. mentella* | *S. mentella (shallow)* |
| **Liverpool_1** | croaker | UK | *S. marinus* (also *S. viviparus*) | *S. norvegicus* (also *S. fasciatus* / *S. mentella*) | *S. norvegicus* |
| **Liverpool_2** | croaker | UK | *S. marinus* (also *S. viviparus*) | *S. norvegicus* (also *S. fasciatus* / *S. mentella*) | *S. norvegicus* |
| **Liverpool_3** | red Snapper | UK | *S. marinus* (also *S. viviparus*) | *S. norvegicus* (also *S. fasciatus* / *S. mentella*) | *S. norvegicus* |
| **Liverpool_4** | red Snapper | UK | *S. marinus* (also *S. viviparus*) | *S. norvegicus* (also *S. fasciatus* / *S. mentella*) | *S. norvegicus* |
| **Liverpool_5** | red Snapper | UK | *S. marinus* (also *S. viviparus*) | *S. norvegicus* (also *S. fasciatus* / *S. mentella*) | *S. norvegicus* |
| **Munich_1** | redfish | Germany | N/A | *S. mentella* | *S. mentella (shallow)* |
| **Munich2** | redfish | Germany | N/A | *S. mentella* | *S. mentella (shallow)* |
| **Munich_3** | redfish | Germany | *Sebastes* sp. | *S. mentella* (also *S. marinus*) | *S. mentella (deep)* |
| **Munich_4** | redfish | Germany | *Sebastes* sp. | *S. marinus* (also *S. mentella*) | *S. mentella (deep)* |
| **Hamburg_1** | *S. norvegicus* | Germany | *Sebastes* sp. | *S. marinus* / *S. mentella* | *S. mentella (shallow)* |
| **Hamburg_2** | *S. alutus* | Germany | S. polyspinis | Unidentified | Unidentified |
| **Hamburg_3** | *S. norvegicus* | Germany | *S. marinus* (also *S. viviparus*) | *S. norvegicus* (also *S. fasciatus* / *S. mentella*) | *S. norvegicus* |
| **Hamburg_4** | *Sebastes sp.* | Germany | *S. marinus* (also *S. viviparus*) | *S. norvegicus* (also *S. fasciatus* / *S. mentella*) | *S. norvegicus* |
| **Leuven_1** | *S. norvegicus* | Belgium | *S. marinus* (also *S. viviparus*) | *S. fasciatus* / *S. marinus* / *S. mentella* | *S. norvegicus* |
| **Leuven_2.1** | *S. norvegicus* | Belgium | *S. marinus* (also *S. viviparus*) | *S. norvegicus* (also *S. fasciatus* / *S. mentella*) | *S. norvegicus* |
| **Leuven_2.2** | *S. norvegicus* | Iceland | *S. marinus* (also *S. viviparus*) | *S. norvegicus* (also *S. fasciatus* / *S. mentella*) | *S. norvegicus* |

N/A: no available sequence; *S. marinus* = *S. norvegicus*

**Table 5.2** Blind samples analysed for species identification with label information and results from BOLD and Genbank databases using the mtDNA COI and control region respectively. Species i.d. is inferred from phylogenetic reconstruction based on d-loop sequences (see Figure 5.2). Classification refers to original MarRef identification (see methods).

| Code | BOLD database (COI) >99% | Genbank (D-loop) >99% | Species i.d. (d-loop) | classification |
|---|---|---|---|---|
| **B_1** | N/A | *S. mentella* | *S. mentella* (deep) | *S. mentella* Icelandic shelf west |
| **B_2** | *S. marinus* (also *S. viviparus*) | *S. norvegicus* (also *S. fasciatus* / *S. mentella*) | *S. norvegicus* | Right |
| **B_3** | *S. viviparus* | *S. viviparus* | *S. viviparus* | Right |
| **B_4** | *S. viviparus* | *S. viviparus* | *S. viviparus* | Right |
| **B_5** | *S. marinus* (also *S. viviparus*) | *S. norvegicus* (also *S. fasciatus* / *S. mentella*) | *S. norvegicus* | Right |
| **B_6** | N/A | *S. mentella* | *S. mentella* (deep) | *S. mentella* (shallow) |
| **B_7** | N/A | *S. mentella* | *S. mentella* (shallow) | *S. mentella* Icelandic shelf west |
| **B_8** | *S. viviparus* (also *S. marinus*) | *S. viviparus* | *S. viviparus* | Right |
| **B_9** | N/A | *S. norvegicus* (also *S. fasciatus* / *S. mentella*) | *S. norvegicus* | Right |
| **B_10** | *Sebastes* sp. | *S. mentella* | *S. mentella* (deep) | Right |
| **B_11** | *Sebastes* sp. | *S. mentella* | *S. mentella* (shallow) | Right |
| **B_12** | *Sebastes* sp. | *S. mentella* | *S. mentella* (shallow) | *S. mentella* Icelandic shelf west |
| **B_13** | *S. viviparus* | *S. viviparus* | *S. viviparus* | Right |
| **B_14** | *Sebastes* sp. | *S. mentella* | *S. mentella* (shallow) | Right |
| **B_15** | S viviparus (also *S. marinus)* | *S. viviparus* | *S. viviparus* | Right |
| **B_16** | N/A | *S. mentella* | *S. mentella* (shallow) | *S. mentella* (deep) |
| **B_17** | N/A | *S. mentella* | *S. mentella* (shallow) | Right |
| **B_18** | *S. mentella* (also *S. marinus*) | *S. mentella* | *S. mentella* (deep) | Right |
| **B_19** | *S. mentella* (also *S. marinus*) | *S. norvegicus* (also *S. fasciatus* / *S. mentella*) | *S. norvegicus* | *S. mentella* (shallow) |
| **B_20** | *Sebastes* sp. | *S. marinus* (also *S. mentella*) | *S. mentella* (shallow) | *S. norvegicus* |
| **B_21** | *S. mentella* (also *S. marinus*) | *S. mentella* | *S. mentella* (deep) | Right |
| **B_22** | *S. mentella* | *S. mentella* | *S. mentella* (deep) | Right |
| **B_23** | *S. mentella* also *S. marinus* | *S. mentella* | *S. mentella* (deep) | *S. mentella* (shallow) |
| **B_24** | *Sebastes* sp. | *S. mentella* | *S. mentella* (shallow) | *S. mentella* (deep) |
| **B_25** | *Sebastes* sp. | *S. mentella* | *S. mentella* (shallow) | Right |
| **B_26** | *S. mentella* (also *S. marinus*) | *S. mentella* | *S. mentella* (deep) | Right |
| **B_27** | *Sebastes* sp. | *S. mentella* | *S. mentella* (shallow) | Right |
| **B_28** | *S. mentella* | *S. mentella* | *S. mentella* (shallow) | *S. mentella* (deep) |
| **B_29** | N/A | *S. mentella* | *S. mentella* (deep) | Right |

N/A: no available sequence; *S. marinus* = *S. norvegicus*

### 5.4.3 Phylogenetic analysis

Phylogenetic reconstruction under a neighbor-joining framework was conducted with references haplotypes for each species, *S. viviparus*, *S. fasciatus*, *S. norvegicus*, and shallow- and deep-pelagic *S. mentella*. Included in the phylogenetic analysis were 29 blind specimens and 19 market samples. Selected Pacific *Sebastes* sequences were included in subsequent analysis because one product was sold as *S. alutus* and therefore set *S. aleutianus* as outgroup (Figure 5.1, 5.2; see Hyde & Vetter, 2007 for accession numbers). Both mitochondrial genes show discordant phylogenetic patterns. The COI topology was ambiguous since *S. fasciatus* and *S. mentella* appear to be paraphyletic, while the d-loop topology was consistent with the specific definition of each taxon, strongly recovering each North Atlantic species in a given cluster, including shallow- and deep-pelagic *S. mentella* (Figure 5.1, 5.2). Thus the criteria used to assess species/population assignment of blind and market specimens were based on inferences from the d-loop phylogeny.

Identification of blind specimens inferred from the d-loop phylogeny was cross-referenced against records from the Marine Research Institute - MarRef. Assessment of blind test samples were generally in agreement with the exception of samples "B_19 and B_20" which cluster with *S. norvegicus* and *S. mentella* (deep) despite being visually classified as *S. mentella* (shallow) and *S. norvegicus* respectively (Table 5.2, Figure 5.2). Furthermore, five mismatches were identified concerning the shallow- and deep-pelagic *S. mentella* types. Two specimens (B_6 and B_23) cluster with the deep-pelagic group that were previously identified as belonging to the shallow-pelagic group. Similarly, three specimens (B_16, B_24 and B_28) cluster with the shallow-pelagic group but were originally identified to belong to the deep-pelagic group.

For the market label analysis, *S. norvegicus* and *S. mentella* (shallow- and deep-pelagic types) were the most commonly available species in markets. Products from Belgium were labelled *S. norvegicus* and nest with the *S. norvegicus* references. German products from Munich sold as 'Rotbarsch' (*Sebastes* spp.) clustered with shallow- (Munich_1 & Munich_2) and deep-pelagic (Munich_3 & Munich_4) *S. mentella*. In Hamburg, however, one sample labelled '*Sebastes* spp.' grouped with *S. norvegicus* (Hamburg_4) while two products sold as '*S. norvegicus*' and '*S. alutus*' group with shallow-pelagic *S. mentella* (Hamburg_1) and Pacific *Sebastes* respectively. Although the sample sold as '*S. alutus'* could not be reliably resolved at species level (Figure 5.2). Norwegian products sold as 'redfisk' (redfish) grouped with *S. norvegicus* (Bergen_1) and

deep-pelagic *S. mentella* (Bergen_2) while one product from Oslo grouped with shallow-pelagic *S. mentella* (Oslo_3). Among the UK specimens, five samples sold as 'croaker' (N=2, Liverpool_1 & Liverpool_2) and 'snapper' (N=3, Liverpool_3-L Liverpool_5) cluster with *S. norvegicus*.

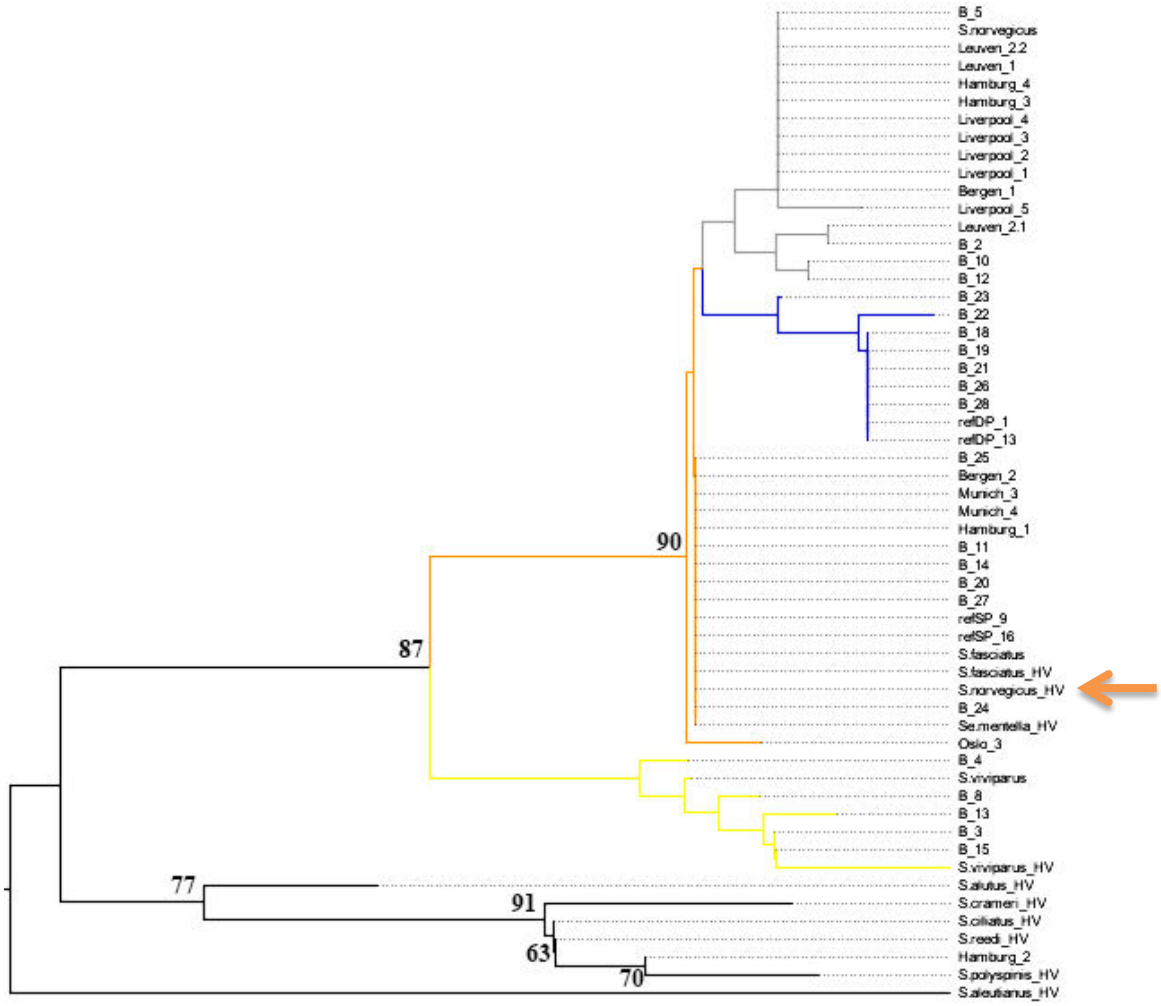



**Figure 5.1** Neighbor-Joining tree of mtDNA COI sequences. Phylogenetic tree using the 7 generated reference sequences representing the four recognized redfish species and *S. mentella* shallow- and deep-types and including COI sequences of 10 species from Hyde and Vetter (2007). Numbers at the nodes represent bootstrap support, with <50% absent after 1,000 replicates. HV indicates Hyde & Vetter's (2007) generated sequences and orange arrow shows erroneous reference entry. Reference entries (this study): *S. viviparus*, *S. fasciatus*, *S. norvegicus* shallow-pelagic *S. mentella*: refSP, deep-pelagic *S. mentella*: refDP). Colour of the branches represent species/population assignment, yellow: *S. viviparus*; gray: *S. norvegicus*; orange: *S. fasciatus*/*S. mentella* (shallow); blue: *S. mentella* (deep).

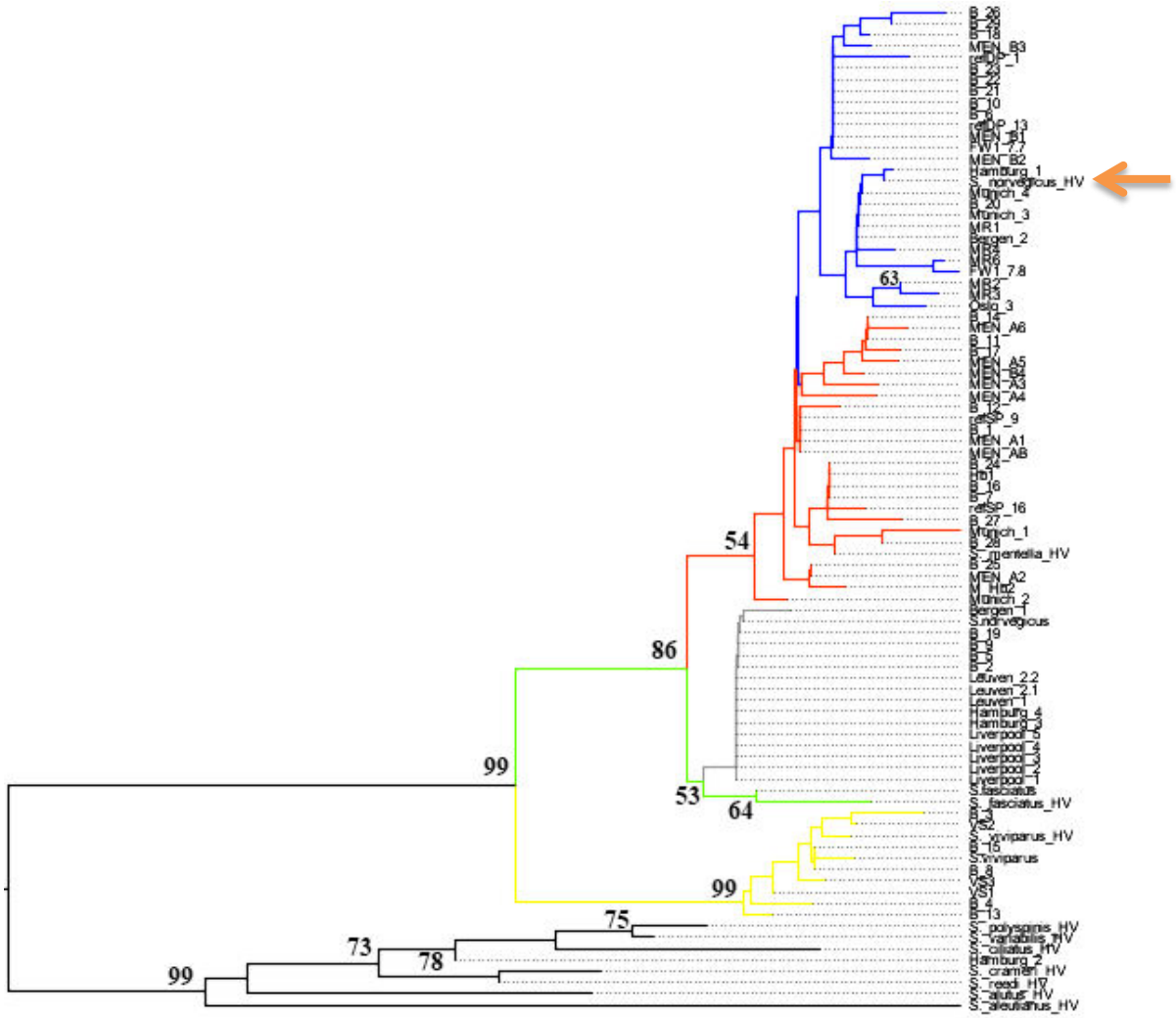



**Figure 5.2** Neighbor-Joining tree of mtDNA control region sequences. Phylogenetic tree using the 7 generated reference sequences representing the four recognized redfish species and *S. mentella* shallow- and deep-types and including control region sequences from10 Pacific species from Hyde and Vetter (2007). Numbers at the nodes represent bootstrap support, with <50% absent after 1,000 replicates. HV indicates Hyde & Vetter's (2007) sequences and orange arrow shows erroneous reference entry. Reference entries (this study): *S. viviparus*, *S. fasciatus*, *S. norvegicus* shallow-pelagic *S. mentella*:refSP, deep-pelagic *S. mentella*: refDP). Colour of the branches represent species/population assignment, yellow: *S. viviparus*; S. fasciatus: green; gray: *S. norvegicus*; red: *S. mentella* (shallow); blue: *S. mentella* (deep), FW: *S. mentella* (deep – Faroe west). From Artamonova *et al*. 2013; VS: *S. viviparus*; MEN: *S. mentella*; M_hb2 & hb1: *S. mentellaS.viviparus* hybrid; MR: *S. norvegicus* (but these haplotypes group with *S. mentella* (deep).

## 5.5 Discussion

North Atlantic *Sebastes* is represented by four species (*S. viviparus*, *S. fasciatus*, *S. norvegicus* and *S. mentella*), that can be sold under a common market name, "redfish", which does not have a uniform designation of species in different European countries. The practice of marketing multiple species under one vernacular name hinders the

identification of sea food products and consequently has important implications for consumer choice and conservation (Griffiths *et al*., 2013). Thus DNA based methods are often required to authenticate the correct labelling of fish, particularly when morphological traits are removed. Here we show that the perception of much of the biocomplexity of *Sebastes* is compromised by the use of the 'redfish' umbrella term. We also find that, irrespective of the market name used, the universal COI approach to identify species is inadequate for this group. The use of a more variable mtDNA fragment (d-loop) allows for the distinction of monophyletic groups, including the four species and the two *S. mentella* "shallow" and "deep" types (Shum *et al*., 2015). However, the use of public data bases is currently affected by wrong reference sequence entries.

### 5.5.1 Complicating issues for DNA barcoding for recently radiating species

North Atlantic redfish species have a recent evolutionary history, having diversified during the Pleistocene (Hyde and Vetter, 2007; Shum *et al*., 2015), and exhibit overlapping meristic and morphological characteristics. This makes phenotypic-based species identification difficult between these closely related species, which may result in misclassification of individuals (Pampoulie & Daníelsdóttir, 2008). Many studies have shown the effectiveness of DNA barcoding using the mitochondrial cytochrome c oxidase I (COI) gene for species identification in a wide range of animal species (Wong & Hanner, 2008; Filonzi *et al*., 2010). However, universal COI barcoding presents a challenge to discriminate between very similar radiating species due to the highly conserved COI gene among congeneric species (Steinke *et al*, 2009; Viñas & Tudela, 2009). Through this study, we discovered many limitations that apply to *Sebastes* DNA-based identification.

Firstly, the BOLD repository contains a limited availability of representative voucher specimens and reference sequences. This is particularly the case for Atlantic redfish as the BOLD search of the COI barcodes identified multiple species that fall within the 2% divergence threshold. These matches included similar identity scores between *S. norvegicus* and *S. viviparus* and between *S. mentella* and *S. norvegicus*. Although the pool of blind specimens did not include *S. fasciatus*, a BOLD query of the reference samples could not be identified at species level (i.e. *Sebastes* sp.). Similarly, the BOLD database failed to identify 8 of the blind specimens to species level that were classified to be *S. mentella*. Thus DNA barcoding has poor resolution for this young group due to insufficient

reference sequences and/or lacks the diagnostic polymorphisms to accurately identify them.

The mitochondrial control region (d-loop) however is a fast evolving marker and has been reported to reliably distinguish between North Atlantic redfish species (Shum *et al.*, 2015). D-loop sequences queried against the Genbank database produced a higher number of positive and unambiguous matches and resolved the unidentified specimens that the COI failed to classify (e.g. Oslo_3, Munich_3, Munich_4, Hamburg_1, B_10, B_11, B_12, B_14, B_20, B_24, B_25 & B_27; Table 5.1). Genbank is a larger database than BOLD and contains more sequences that will subsequently increase the identification of unknown samples. However, it contains a mixture of verified and unverified sequences without appropriate quality-control procedures which may populate the database with ambiguous sequence submissions (Wong & Hanner, 2008). Similar to COI, a BLAST search of the d-loop sequences identified multiple species that fall within the 2% divergence threshold, notably between *S. norvegicus* (100%), *S. fasciatus* (99%) and *S. mentella* (99%) (Table 5.1). Sequences uploaded into Genbank are potentially erroneous due to morphological identification (see below), however, Shum *et al.* (2015) provide a comprehensive dataset of all North Atlantic *Sebastes* (d-loop) in Genbank (see Chapter 2), which can be used as a good source of reference.

While the d-loop exhibits the resolution to unambiguously distinguish among the four species and the two main lineages within *S. mentella* (Figure 5.2), the systematic bias of erroneous/misassigned sequences in the database hinders the process of identification. Furthermore, it should be highlighted that the use of mitochondrial genes is limited because only the matrilineal lineage is examined, which can limit the correct identification of a species, in those groups known to hybridise. Pampoulie & Daníelsdóttir (2008) reported evidence of introgressive hybridization among North Atlantic redfish and discovered individual misclassification based on morphological identification. Consequently, hybrid redfish will be assigned to the maternal parent, which can further complicate the application of molecular-based identification approaches (Nicolè *et al.*, 2012).

### 5.5.2 Phylogenetic representation

The phylogenetic approach was based on the construction of a Neighbor-joining tree to assess the results of the database searches (Figure 5.1, 5.2). The NJ reconstruction of the

COI marker was able to distinguish between *S. viviparus*, *S. norvegicus* and the deep-pelagic *S. mentella*, but was not diagnostic between *S. fasciatus* and shallow-pelagic *S. mentella*. However, the phylogenetic reconstruction of the mtDNA control region allowed full discrimination of all species into distinct monophyletic groups, including population level assignment of *S. mentella* (Figure 5.2). The assessment of the blind specimens revealed unambiguous clustering of species with the exception of two misclassifications (B19 & B20). These specimens were morphologically classified by visual eye but were identified as another species using the mtDNA control region (Table 5.2). This, perhaps, is indicative of misidentification or potentially hybridisation (Pampoulie & Daníelsdóttir, 2008). The addition of neutral nuclear markers is required to determine the presence of hybrids. Moreover, we found discrepancies concerning the visual identification of the shallow- *vs.* deep-type *S. mentella.* These two types are often separated by their morphological appearance (Stefánsson *et al*., 2009a), their depth and geographic distribution, yet 17% (5 of 29) of the blind specimens were inaccurately identified. These patterns may have cascading consequences for conservation management and seafood labelling.

The market-level analysis exemplifies patterns of misidentification for North Atlantic redfish. Among the market collections, two North Atlantic (*S. norvegicus* and *S. mentella* shallow- and deep-types) and one ambiguous North Pacific species were recorded. Redfish collected in the UK were marketed under various, inaccurate labels, including 'croaker' and 'red snapper', all of which were identified as *S. norvegicus*. This is an alarming discovery and it demonstrates a severe lack of knowledge, raising serious concern about the redfish industry in the UK. In Europe it is required that the names of fish products contain the scientific name to be labelled on products (EC No 1379/2013). However, in some cases, German fish mongers or companies write "Rotbarsch or *Sebastes* spp." because the German index of designation names for fish species allow species in the genus *Sebastes* spp. not identified to species level may be called "Rotbarsch" (Federal Agency for Agriculture and Food, 2015). Therefore, products from Munich and one product from Hamburg fall under the umbrella term "Rotbarsch". While the majority of the labelled redfish products were grouped with the species listed on the label, there were instances of mislabelling. Two samples, Hamburg_1 and Hamburg_2, exhibit mitochondrial sequences that did not group with the designated label product indicating that these products are mislabelled. In Norway, the designation name for fish species sold

as "redfisk" can only refer to *S. marinus* (*S. norvegicus*), while *S. mentella* must be correctly labelled *S. mentella* (Ministry of Industry and Ministry of Fisheries, 2013). Therefore, incidence of mislabelling were recovered from Bergen (N=1) and Oslo (N=1) as these products were labelled "redfisk" but identified as *S. mentella*.

An important issue revealed in the present study concerned the sale of North Atlantic *Sebastes* under a single vernacular name, redfish. In Europe, the umbrella term "redfish" show conflicting definitions among nations and thus creates confusion in the redfish industry regarding which species are targeted under this term. The practice of selling multiple species under a common name has important consequences for market driven conservation strategies and prevents consumers from making informed decisions about sustainable purchasing (Logan *et al*., 2008). For example, two "redfish" samples collected from one retailer in Bergen, Norway, clustered as two different species, *S. norvegicus* and *S. mentella*, which have markedly different life history characteristics and conservation needs (Table 1; Cadrin *et al*., 2010; Marine research Institute, 2014). Moreover, *S. mentella* possess at least two genetically distinguishable groups with different management advice, with a moratorium placed on the shallow population and the deep population limited to 20,000 tonnes, (Cadrin *et al*., 2010; Shum *et al*., 2015) yet, we identified both *S. mentella* types in Germany under the common name, Rotbarsch (redfish), which eliminates the ability of consumers to choose less vulnerable groups.

### 5.5.3 Resolution of species-population identity

An important pattern revealed in the present study concerns the inclusion of misclassified redfish species in genetic data bases, which leads to erroneous conclusions when querying sequence repositories. Hyde and Vetter (2007) examined the evolutionary relationships of 101 *Sebastes* species, within which context the four Atlantic species occupy a relatively peripheral position. However, by comparing the topology of the Atlantic clade in their study and the present work, using the same d-loop fragment, we noticed that the reciprocal relationships among *S. mentella*, *S. norvegicus* and *S. fasciatus* were different (see Figure 5.3), with *S. norvegicus* appearing closer to *S. mentella* in Hyde & Vetter (2007), while clustering as sister taxon of *S. fasciatus* in our tree, in line with Shum *et al*. (2015).

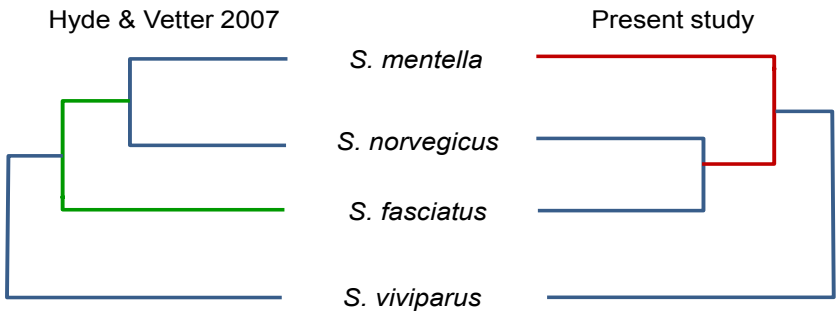



**Figure 5.3** Proposed North Atlantic *Sebastes* phylogenies of Hyde & Vetter (2007) compared to the present study.

Although the topology could be affected by different amounts of sequences compared, we included the sequence data from Hyde and Vetter (2007) and discovered that their *S. norvegicus* sample clustered within the deep-type *S. mentella* (Figure 5.2), indicating that their reference material was a result of a misidentified deep-type *S. mentella* or a *S. norvegicus* X *S. mentella* deep-type hybrid. The distinction between *S. mentella* and *S. norvegicus* has been reported to be extremely difficult using morphological traits (Power & Ni, 1985; Rubec *et al.*, 1991) and it is expected that misclassification can occur. An even more notable example is offered by the comparison with Artamonova *et al.* (2013), who evaluated hybridization of *S. mentella* and reportedly included representative NA redfish species in their study *"which could be identified with certainty by morphology studies and allozyme analysis"* (pp. 1794, paragraph 12 in Artamonova *et al.*, 2013). Strikingly, six of the putative *S. norvegicus* haplotypes representing 15 individuals (collected at Bear Island Trough & Kopytov area, north-east of the Norwegian Sea), all consistently cluster within the shallow- and deep-type *S. mentella* (Figure 5.2). Their *S. norvegicus* haplotypes mostly cluster with the deep-pelagic group from reference samples collected west of the Faroes. Moreover, they seem to maintain integrity from the deep-pelagic group found in the Irminger Sea and west of the Faroes, indicating further population structuring than previously thought (See Figure 5.4, Shum *et al.*, 2015). This example stresses the inherent difficulty in discriminating phenotypes and hence, interpreting results for North Atlantic *Sebastes*. In fact, considering these misidentified deep-pelagic haplotypes in the mtDNA d-loop reconstruction, it seem that many of the *S. mentella* products collected from Germany and Norway come from this extremely understudied stock in the northern Norwegian Seas that seems to consistently fill these markets (Figure 5.2). Thus it is important that publicly available databases are carefully

checked and verified, especially for species that are both "difficult" and important, so that the power of DNA identification can be properly harnessed.

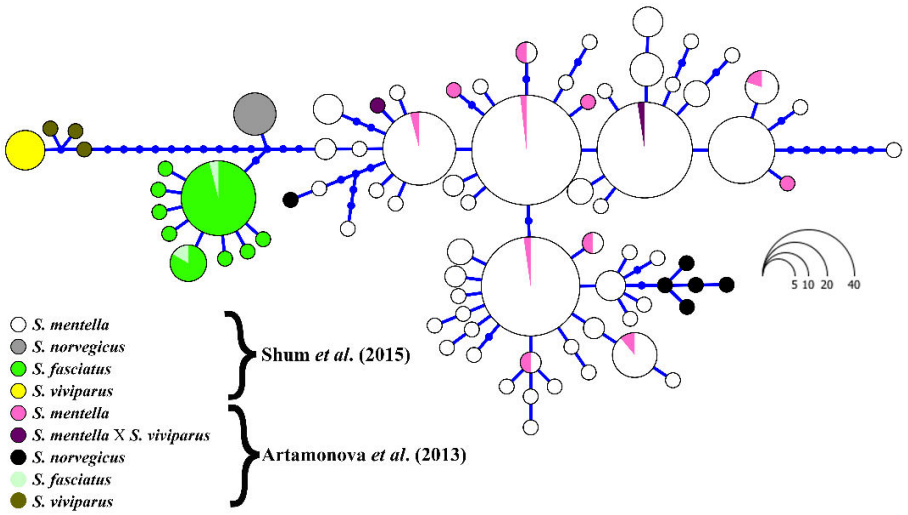



**Figure 5.4** Haplotype genealogy of North Atlantic *Sebastes* from Shum *et al.* (2015). Data from Artamonova *et al.* (2013) is visualised among reference haplotypes where their *S. norvegicus* clusters with shallow- and deep-type *S. mentella*.

## 5.6 Conclusions

Our study focused on the molecular identification of North Atlantic *Sebastes* in order to clarify the status of this young group and present the first attempt to investigate the patterns of sale of Atlantic *Sebastes* across Europe. We confirmed the inadequacy of the COI universal barcode to discriminate among species within this commercially important genus. We also illustrate the difficulty of publicly available databases to provide reliable matches of unknown specimens caused by scarce or erroneous reference entries. However, the use of the d-loop locus in a phylogenetic framework provided a higher resolution in identifying unknown specimens to species/population level and allow traceability of a potentially newly discovered stock north-east of the Norwegian Seas being used to meet market demand. We report examples whereby morphological misidentification may have led to misleading data. Thus a thorough revision of the phylogenetic and biological relationships among these closely related species is required to understand their evolution, if we are to continue harvesting this important commercial resource.

# Chapter VI

# General Discussion

## 6.1 Main findings

The process of speciation is poorly understood in the marine pelagic environment, especially in the absence of obvious geographical barriers. Understanding the effects of geographical barriers on the genetic structure and diversity of marine populations can elucidate the causes that influence diversification, and ultimately, speciation. While ecological speciation can occur between strictly allopatric populations, divergent natural selection can also drive speciation in the face of gene flow (Nosil, 2012). The genus *Sebastes* exhibits a considerable diversity of forms in the absence of geographical barriers and has sparked great interest in the factors promoting speciation in this group (Hyde & Vetter, 2007; Hyde *et al*., 2008). The species subject of this thesis, *Sebastes mentella*, has received substantial interest over the past few decades (Magnússon and Magnússon, 1995; Cadrin *et al*., 2010; Planque *et al*., 2013). The existence of depth segregated *S. mentella* populations in the North Atlantic has triggered extensive morphological and genetic investigations (Cadrin *et al*., 2010). However, the complex issue of depth has hindered a complete understanding of depth-associated stock structure of *S. mentella* and appropriate spatial units for fisheries management, which continue to be hotly debated (Makhrov *et al*., 2011; Cadrin *et al*., 2011).

One major advancement of the present work was to carry out a rigorous sampling design to collect samples at three discrete depths, in several geographic locations, in order to tackle specifically depth-related population structure. This allowed an investigation of the biology of *S. mentella*, keeping the issue of depth always explicitly at the core of the matter. More specifically, in **chapter II** a multiple molecular marker approach was employed to investigate the number of genetically-distinguishable groups of *S. mentella* in the North Atlantic and how they are biogeographically structured and connected; in **chapter III** an examination of spectral absorbance on rod visual pigments and genetic analysis of the rhodopsin gene was assessed in order to link genotype and phenotype for fish collected at different depths and hence, light environments; in **chapter IV** *de novo* transcriptome sequencing (RNA-seq) was carried out to identify depth-associated

candidate genes particularly involved in visual perception and lastly, **chapter V** explores the end point of the North Atlantic *Sebastes* fisheries by carrying out a market product accuracy survey from retailers across Europe. Collectively, these results significantly contribute to the understanding of redfish evolution in the Atlantic as well as the development of effective management and conservation strategies.

In **chapter II**, the first extensive sampling and phylogeographic investigation of pelagic *S. mentella* in the North Atlantic is presented using molecular genetic markers to resolve the interplay between longitudinal, latitudinal and vertical gradients in shaping the structure of this oceanic species. Despite the high commercial value of this species in Europe, little was known about the historical processes and evolutionary trajectories of *S. mentella* populations that make up their current three-dimensional distribution. This study revealed the existence of a shallow-pelagic group (<500m) showing strong homogeneity across the North Atlantic from the Norwegian Sea, east of the Faroes to west and south of the Irminger Sea above 500m, whereas the deep-pelagic group (>500m) were more localised to the central North Atlantic below 500m. Analysis of the deep-pelagic group revealed further contemporary population structure between collections from North-east Irminger Sea and west of the Faroes indicating, perhaps, a role of bathymetric features (such as mid-ocean ridges and slopes) in facilitating the retention and establishment of reproductive isolation by depth. Taken together, mtDNA reflect evolutionary history throughout the late Pleistocene, whereas the integration with microsatellite data serves to better reflect postglacial divergence and the patterns of contemporary gene flow among populations. Figure 6.1 illustrates a possible scenario of the historical movement and evolutionary trajectory of *S. mentella* in the North Atlantic during the Last Glacial Maximum approximately 20,000 years ago. **(a)** Expanding ice sheets may have forced *S. mentella* to southerly latitudes as far as Grand Banks (40˚N) to greater depths where productivity would have been higher during the LGM as global sea level fell by ~120-135m below current levels. **(b)** As the climate warmed, glacial sea ice retreated and *S. mentella* had an opportunity to advance in a northerly direction south of Iceland ~13 kyr. **(c)** However, as the glaciers collapsed initiating the end of the last ice age, intense calving in the northern hemisphere ice sheets during the Younger Dryas (YD, 13-11 kyr BP) resulted in massive icebergs advancing in a southerly direction forcing the North Atlantic into near-glacial conditions (Crucifix & Berger, 2002). The onset of the YD may have favoured allopatric conditions between the two groups as the deep-pelagic group occupied deeper refuges while **(d)** the shallow-pelagic rapidly spread with the retreating sea ice and

rising sea level following the YD into the Norwegian and Barents Seas. **(e)** The reformation of the Irminger Sea and complex oceanic conditions at depth allowed for the establishment of contemporary population structure between deep-pelagic groups south-west of the Irminger Sea and west of the Faroes. Finally **(f)** illustrates the currently known distribution of *S. mentella* in the North Atlantic as well as potentially further groups (see Chapter V).

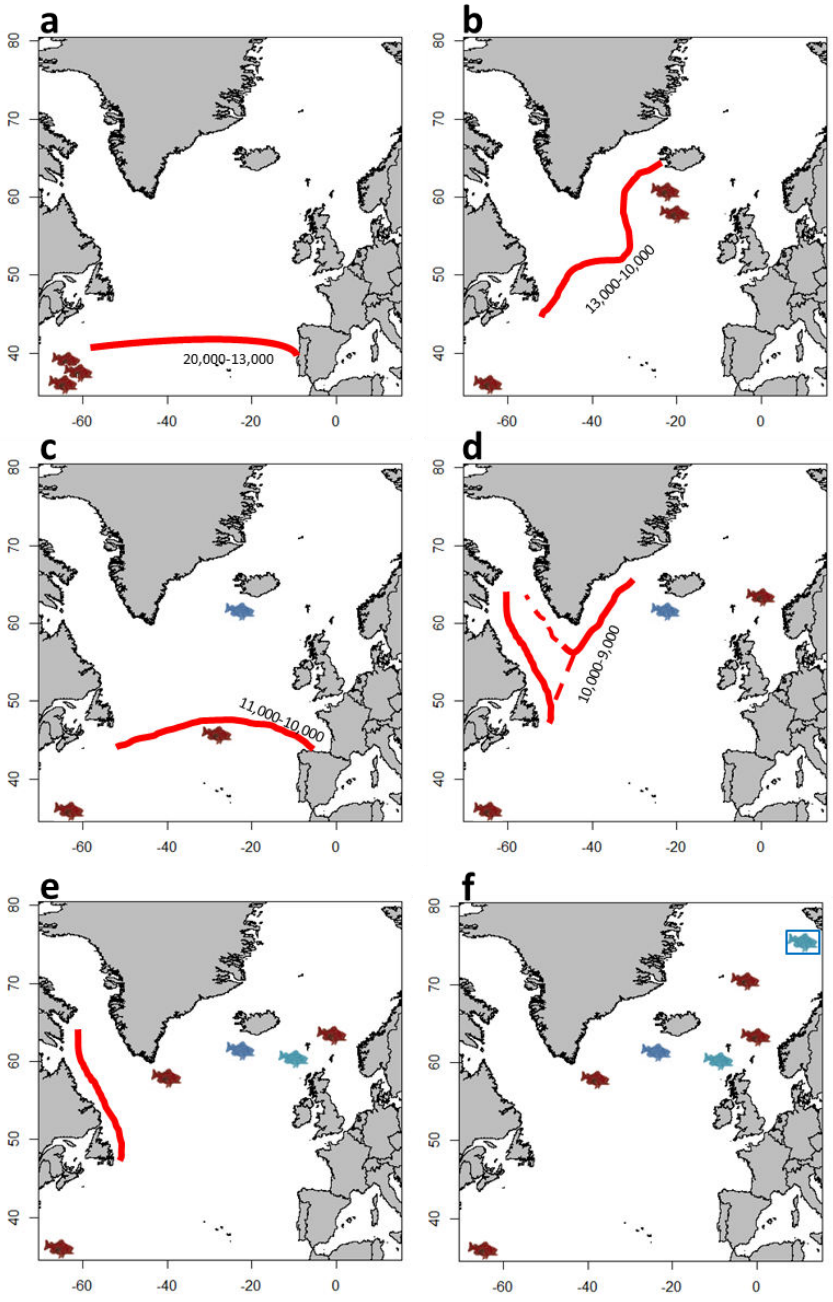



**Figure 6.1** illustrates a plausible scenario of the historical movement and evolutionary trajectory of *Sebastes mentella* during the Pleistocene glaciations. Lines in red represent the limit of expanding sea ice and values indicate years before present. In panel (d) dotted lines indicate sea ice retreat during 10,000-9,000 years ago, (f) indicates area of possible further deep-pelagic sub-structuring (see chapter V). See text for description of each panel.

The historical process and mechanism of isolation as a result of the Pleistocene glaciations raise the possibility of incipient speciation in *S. mentella*. The shallow and deep groups maintain a striking overall integrity despite reported cases of hybridisation (Pampoulie & Danielsdottir, 2008; Shum *et al*., 2015). Furthermore, the significance of these events mirror similar patterns of divergence observed from Pacific *Sebastes*. Typically, colonisation occurs by a stepwise invasion of a newly available habitat in which vicariant isolation results due to local isolating mechanisms such as glacial advance and retreat, sea level change, and ocean currents (Hyde & Vetter, 2007). This diversification by depth, along with the 'younger' age of North Atlantic *Sebastes* compared to the Pacific radiation, is consistent with the early stages of speciation.

*Sebastes* spp. exhibits a wide range of depths from near-surface waters extending down to 1000m (Love *et al*., 2002), and Sivasundar & Palumbi, (2010) first showed that the gene encoding the visual pigment, rhodopsin, is adaptively evolving in response to different depths and hence, light environments. *S. mentella* inhabit varying light levels of the mesopelagic zone and exhibit an amino acid difference at site 119 that strongly discriminates the shallow- and deep-pelagic groups. Therefore in **chapter III** an examination of the rod visual pigments involved in dim-light vision was carried out in order to determine a relationship between genotype and phenotype. The beaked redfish offers an interesting study system to examine spectral sensitivity of rhodopsin because they possess unique depth distributions that expose the shallow and deep-pelagic groups to different light intensities. However, data collected fail to capture the paradigm of adaptive tuning of visual pigments by amino acid changes associated with depth, and point towards a role in a physiological response involving the duration of the photoresponse and daily vertical migrations between the two groups.

In **chapter IV**, the *de novo* assembly of RNA-seq data of *Sebastes* retinal transcriptome is presented. This enabled an examination of transcriptomes without the need for a reference genome sequence. RNA-seq data identified the presence of functional opsin genes involved in colour perception and physiological processes that control several aspects of photoresponse termination, and remarkably, revealed an association of depth-related immunity genes between the shallow- and deep-pelagic groups. The significance of these findings is threefold and has powerful resonance for understanding adaptive evolution of *Sebastes*.

First, the deep-pelagic group experience higher expression levels of the G protein-coupled receptor kinase 4 (GRK4), although not the main protein involved in rhodopsin

kinase, has previously been shown to phosphorylate rhodopsin in vitro. This suggests the shallow group have greater rhodopsin sensitivity and a longer photoresponse compared to the deep group. This response perhaps evolved as a consequence of the shallow group expanding into the North Atlantic to greater ambient light levels during rising sea levels. However, further study is required to make an association between photoresponse duration and the rate of phosphorylation in the outer segment of the rods for this group (Arshavsky, 2002). Second, *S. mentella* show the presence of multiple opsin genes that show different patterns of expression. Both groups possess low levels of expression in violet- and ultraviolet-sensitive opsins suggesting they were "non-functional" at the time of capture, while blue- and green-sensitive opsins were highly expressed above housekeeping gene elongation factor-1α. It is unclear why *S. mentella*, a deep-sea species, would require the use of colour perception in such dimly lit environments. Nevertheless, daily diel migrations and the consumption of bioluminescent prey (Douglas & Partridge, 1997; Turner *et al*., 2009) may explain these patterns as they become exposed to wavelengths in the blue-green spectra of light, although the extent of bioluminescence exposure between groups remains unclear. Furthermore, the presence of multiple opsin genes may reflect phylogenetic constraints of ancestral genes in the genomic evolution of *Sebastes*. The *Sebastes* radiation has been compared to the more recent radiation of cichlids in the Great African lakes (Greenwood, 1991), in which cichlids are notable for their diversity in sexual dimorphism in colour patterns and visual communication processes governing mate choice (Jordan *et al*., 2003; Kocher, 2004). The cones allow an individual to distinguish contrast and colour perception which likely provides a significant advantage as *S. mentella* invade new environments composed of bioluminescent prey that make them successful visual predators in the deep sea. Thus the role of colour vision, perhaps, influences mate choice, which could act as a barrier for reproductive isolation in North Atlantic *S. mentella*. Ideally, observational behavioural studies of reproductive strategies in Atlantic *Sebastes* would be of great value in this context.

Lastly, immune-related genes were significantly overexpressed in the deep-pelagic group and two explanations may serve to elucidate the patterns observed. First, the immune genes are ubiquitously expressed in all tissues. Previous studies have documented an association of MHC variation and parasitic load (Wegner *et al*., 2003), and biological tags such as abnormalities in pigmentation and parasite infestation are routinely used to discriminate *S. mentella* types. The shallow-pelagic group experience significantly higher rates of parasite infestation (reviewed in Cadrin *et al*., 2010) while the deep-pelagic

group's greater resistance to infection may well be linked to the greater expression of immune genes. Second, the immune genes show tissue-specific gene expression in the retina. This suggests that visual systems of the deep-pelagic group may have increased protection from infection to avoid visual impairment given the light-induced stress in darker environments.

Overall, RNA-seq was an attempt to identify depth-associated candidate genes involved in the perception of light. Analyses of retinal transcriptomes highlight *S. mentella* to be an extremely successful lineage, and owe their successful adaptation to phylogenetic constraints which may have allowed the colonisation to different depth environments through the exploitation of colour perception and immunity.

The final study (**chapter V**) addresses the status of redfish sold in Europe by evaluating the performance of molecular markers against publicly available databases. The morphological identification of North Atlantic redfish remains controversial and misidentification may likely produce misleading data, with cascading consequences that negatively affect conservation and seafood labelling. This study examined the strength of mtDNA genetic markers to identify North Atlantic *Sebastes* using "blind" specimens and was subsequently applied to market specimens collected from four European countries. Results show the inadequacy of the COI universal barcode to discriminate among species, while phylogenetic analysis based on mtDNA control region sequences provided a higher resolution for species identification. Furthermore, public databases revealed the existence of inaccurate reference sequences, likely stemming from species misidentification from previous studies, which currently hinder the efficacy of DNA methods for the identification of *Sebastes* market samples. However, in order to fully comprehend the status of each specimen, the use of nuclear markers is required to identify potential hybrids. An important discovery originating from inaccurate reference entries is the detection of further population sub-structure among deep-pelagic *S. mentella* in the Norwegian Sea, in which Artamonova *et al.* (2013) incorrectly identified these samples as *S. norvegicus*. This finding highlights a fundamental gap in knowledge that can strongly impact the assessment, management and viability of fish stocks. Thus a thorough evaluation of North Atlantic *Sebastes* is required if we are to continue harvesting this important commercial fishery resource.

## 6.2 Divergence along an environmental depth gradient

In recent decades there has been considerable progress in identifying the scales, mechanisms and forces driving the formation of species (Butlin *et al.*, 2012; Bowen *et al.*, 2013). However, little attention focused on the mechanisms influencing deep-sea ecosystems that trigger a demersal or pelagic behaviour. The lack of clear isolating barriers and the scale at which divergence can occur among conspecifics indicate that selection influences depth-related divergence. A number of environmental factors considerably change with increasing depth, including temperature, salinity, oxygen, hydrostatic pressure and light, which independently or in combination, may lead to divergence, and are frequently invoked in mediating adaptation (Somero, 1992; Brown & Thatje, 2011), promoting population divergence and speciation (Rogers, 2000; Bongaerts *et al.*, 2011). Identifying the environmental forces that shape depth-segregated genetic variation will ultimately depend on the system in question and requires further investigation. However, geographic patterns in population structure provide several lines of evidence for identifying the selective forces and habitat boundaries that might isolate or constrain gene pools in marine ecosystems.

Consistent with depth and its associated factors playing an important role in diversification of marine species, several studies have documented strong bathymetric divergence suggestive of cryptic species in Pacific *Sebastes* (Hyde *et al.*, 2008). In addition, divergence is found to be greater among *Sebastes* separated vertically than horizontally (Ingram, 2010). *S. mentella* provide a powerful and compelling case of depth-related ecological divergence influenced by past climactic events, in which, barriers to gene flow evolved between populations as a result of ecologically based divergent selection between depth environments (Nosil, 2012). *S. mentella* likely exhibited a "habitat-first" mode of divergence (see Ackerly, Schwilk & Webb, 2006) during which redfish became initially isolated by depth (e.g. in allopatry) created by sea level fluctuations during interglacial periods of the Pleistocene. This spatial separation likely resulted in individuals reproducing, thereby establishing populations in separated depth habitats. This is evident by the concordance between neutral genetic markers seen from mitochondrial and nuclear markers (Chapter II). Subsequently, variation in traits that affect fitness in different depth habitats resulted in divergent natural selection between shallow- and deep-pelagic *S. mentella* depth ranges. This is apparent because both groups maintain

a striking overall integrity, despite reported cases of hybridisation (Pampoulie & Danielsdottir 2008; Shum *et al*., 2015).

The diversification of *S. mentella* in the North Atlantic may reflect ancestral constraints in genome evolution. RNA-seq revealed the presence of functional opsin genes that indicate the potential to test new hypotheses. Like cichlids, *Sebastes* appears to diverge in light environments and this body of work has demonstrated the potential of *Sebastes* to undergo adaptive evolution of visual pigment genes along depth gradients. Detailed studies demonstrate the role of sensory drive in cichlid speciation involving complex sexual signals and preference between environments (Malinsky *et al*., 2015). Thus, perhaps *Sebastes* follow a similar mechanism in the marine environment given the diversity of forms among species. Similar to cichlids, *Sebastes* is an attractive case to study mechanisms of speciation and sensory drive in marine adaptive radiations, although the long generation time and the difficulty of accessing and observing live specimens will always remain a significant obstacle.

## 6.3 Implications for fisheries and biodiversity

*Sebastes* spp. are ovoviviparous (i.e. internal fertilisation), long-lived, slow-growing, late-maturing, and have low natural mortality. These life-history characteristics influence complex population structure (Cadrin *et al*., 2010; Shum *et al*., 2015) and make populations highly vulnerable to fishing and slow to recover when depleted (Planque *et al*., 2013). Furthermore assessing the health of *S. mentella* stocks is hampered by difficulties in adequate monitoring that include the extent of geographical and depth distribution of *S. mentella*, occupation of demersal and benthic habitats at different life cycle stages, unknown recruitment control, species identification and stock identification and generally poor knowledge of *S. mentella* interactions with the pelagic ecosystems of the Irminger and Norwegian Seas (Planque *et al*., 2013). Therefore, monitoring *S. mentella* requires strong international support and sustained coordination effort by transnational organisations, such as ICES.

The aim of this present study was to map the extent of geographic population structure and depth distribution of *S. mentella* in the North Atlantic. The multi-sampler proved to have a significant advantage over traditional methods because genetic analyses clearly define redfish to occur in areas of overlap. Annual multispecies research trawl surveys in Canadian waters do not separate redfish species as they are difficult to separate

by anatomical characters and it has been noted that managing and monitoring the two stocks by depth is impractical, despite reported cases of depth-related structure (Valentin *et al*., 2014). Management units take the form of geographic proxies, whose boundaries are based on the spatial distribution pattern of the fishery. However, it can be expected that *Sebastes* are sensitive to climate-related range shifts (Hyde & Vetter, 2007, Shum *et al*., 2015), and thus the importance of monitoring specific depths not only minimizes mixed stock catches but may also help forcasting the trajectory of species' ranges (Perry *et al*., 2005; Lenoir & Svenning, 2015).

Genetic analyses demonstrate the existence of at least two genetically distinguishable groups in the Irminger Sea and adjacent waters and refute any notion of panmixia. In fact, analyses revealed that the shallow-pelagic group show strong homogeneity throughout the North Atlantic whereas the deep-pelagic group form rather distinct 'pockets', suggesting their evolutionary distinction also involves different life histories, as shallow- pelagic redfish appear more prone to connectivity. Furthermore, while it was previously thought that the deep-pelagic redfish were more strictly associated with their local habitats in the central north Atlantic (Shum *et al*., 2015), recent evidence point to further sub-structuring in the deep-pelagic lineage (see chapter V). This finding stresses the lack of knowledge surrounding the North Atlantic *Sebastes* which hinders the efficacy of DNA methods for the identification of seafood. North Atlantic *Sebastes* support an important fishery industry, therefore, every effort should be made to monitor and sample *S. mentella* as it is not clear whether these misidentified samples form distinct populations.

The complex life cycle of redfish is generally poorly understood including the reproductive behaviour of *S. mentella* that make them vulnerable to even low levels of overharvesting (Planque *et al*., 2013). Such life cycles may involve extensive migrations, varying distributions of life stages or use of specific habitats such as nursery and feeding grounds. *S. mentella* is an important biodiversity component and possesses adaptive features that make it a key player in the North Atlantic meso-/bathy-pelagic habitat. Therefore, it is important to take the appropriate steps (by targeting specific depths where high density redfish assemblages occur in the North Atlantic) to protect *S. mentella* in the long term, for its continual exploitation as a resource and to conserve the genetic resources for evolutionary potential.

## 6.4 Future directions

The overarching theme throughout this thesis has been to investigate the role of depth in shaping genetic and phenotypic patterns of variation in the beaked redfish, *Sebastes mentella*, in the North Atlantic. The present body of work underpins the usefulness of the methods employed and highlights their potential use for species/stock identification and catapults our current understanding of depth-associated divergence of marine adaptive radiations of *Sebastes*. Although a variety of questions have been addressed, several new questions emerged and require further work in order to fully realise the extent of divergence. Therefore, a number of areas of further study which incorporate many aspects of the above chapters can also be identified by looking at the system as a whole.

Failure to recognise distinct populations within a rapidly divergent lineage will mask population declines and the risk of extinction – which potentially leads to the elimination of incipient species – may be higher than previously believed. This project generated a large tissue bank of reference material of North Atlantic *Sebastes*, including advanced sampling of both *S. mentella* types. This provides a necessary opportunity to obtain a more comprehensive picture of North Atlantic redfish in order to 1) re-evaluate extent of *S. mentella* population structure in the wider North Atlantic by targeted sampling in Canadian waters and the Norwegian and Barents Seas and 2) sample juveniles in nursery areas during periods of larval extrusion to determine the genetic makeup between juveniles and adult stocks. This will clarify whether there is any spatial or temporal segregation/ mixture among different stocks, or even species, during early life stages.

Depth plays an important role in adaptation and speciation. RNA-seq of *S. mentella* retinas revealed significant functional diversity and revealed candidate genes involved in vision and immunity. Firstly, retinas from redfish are needed to determine a functional opsin phenotype by morphological staining and direct measurement of opsin pigments by microspectrophotometry (MSP). Furthermore, gene expression of each opsin is required to assess sensitivity between groups, and measuring the bioluminescence wavelengths of their prey can provide a relationship between vision and diet. Secondly, sampling juvenile retinas may determine a functional switch in opsin genes. By quantifying opsin transcripts by RT-qPCR on retinal tissue of different opsins allows an indirect method to study behaviour. This allows testing whether redfish utilise ultraviolet or violet wavelengths as juveniles or clarify if these opsins have been "lost" during adaptation to dim-light environments in the deep-sea. Lastly, this project revealed a correlation of immune-related

genes associated with depth. Further investigations will be needed to determine the degree of protection the deep-pelagic group experience by examining gene expression patterns from liver and kidney tissue.

To further investigate adaptive divergence and hybridisation the use of Restriction site associated DNA sequencing (RAD-seq) can be implemented to locate genomic regions of divergence that would help pinpointing 'speciation genes' among North Atlantic redfish species. The principal advantage of RAD-seq is the ability to simultaneously examine tens of thousands of genetic loci with vastly reduced sequencing costs versus whole genome approaches. With the availability of a reference genome, this may facilitate in sampling the genomes of multiple individuals to locate "genomic islands" that facilitate speciation.

While North Atlantic redfish are harvested beyond the levels set from scientific advice (Marine Research Institute, 2014), necessary measures are required to determine the frequency and number of species being marketed. Thus a significant effort is required to sample redfish marketed in Europe using methods developed in chapter V. This will allow correct identification of *Sebastes* on a species and population level and determine whether hybrids form a significant percentage of the redfish market.

## 6.5 Final conclusions

*S. mentella* provide an attractive system to study depth-related patterns of substructure. Previous research into North Atlantic redfish returned conflicting evidence of the impact of differentiation amongst regional redfish populations. This thesis set out to understand the mechanisms that shape population divergence along a depth gradient and to evaluate the selective forces that explain their sub-structure. An interdisciplinary approach was carried out using state-of-the-art sampling gear which contributed to the most advanced sampling campaign conducted for *S. mentella* in the Irminger Sea. A suite of molecular methods identified the extent of *S. mentella* biogeography throughout the North Atlantic and by depth, while an extensive examination of their rod visual pigments and retinal transcriptomes provided remarkable insight into adaptation to deep sea environments. Notably, *S. mentella* potentially exploit colour vision in the deep-sea, possibly owing to their success as predatory animals. In addition, the interesting pattern in immune-related processes appears to correlate with differences in depth preference which requires further attention. Less than a decade ago, *S. mentella* was assumed to be panmictic in the North Atlantic and continue to be hotly debated. It is hoped that this body of work will resonate

positively among groups and stakeholders interested in managing this unique biodiversity component, and will succeed in thrusting this research subject beyond obsolete disputes and towards greater achievements and breakthroughs.

# Supplementary Information

## Chapter II

## Three-Dimensional Post-Glacial Expansion and Diversification of an Exploited Oceanic Fish

**Table S0.1** Microsatellite estimates of pairwise genetic differentiation among nine *S. mentella* collections. Pairwise $F_{ST}$ below diagonal and $R_{ST}$ above, based on 10,000 permutations. Significance tested after the False Discovery Rate (FDR) correction.

| | **IrNE$_{DP}$06** | **IrSW$_{SP}$06** | **FW1$_{DP}$07** | **FW2$_{DP}$07** | **FE1$_{SP}$07** | **FE2$_{SP}$07** | **NS$_{SP}$07** | **NIW1$_{SP}$06** | **NIW2$_{SP}$06** |
|---|---|---|---|---|---|---|---|---|---|
| **IrNE$_{DP}$06** | - | **0.068** | -0.044 | 0.009 | 0.106 | 0.034 | **0.094** | **0.236***** | **0.558***** |
| **IrSW$_{SP}$06** | **0.030***** | - | 0.020 | **0.075** | **0.115**** | **0.150*** | **0.153***** | **0.666***** | **0.690***** |
| **FW1$_{DP}$07** | -0.008 | **0.021** | - | -0.062 | **0.066** | -0.014 | -0.008 | **0.169** | **0.267**** |
| **FW2$_{DP}$07** | -0.004 | **0.022**** | -0.009 | - | **0.025** | -0.079 | -0.063 | 0.015 | **0.192**** |
| **FE1$_{SP}$07** | **0.029**** | 0.002 | **0.029**** | **0.022** | - | **0.021** | 0.029 | **0.240** | **0.272**** |
| **FE2$_{SP}$07** | **0.021**** | 0.012 | 0.017 | 0.019 | 0.007 | - | -0.062 | 0.012 | **0.172**** |
| **NS$_{SP}$07** | **0.029***** | -0.003 | **0.024** | 0.010 | -0.003 | 0.011 | - | -0.044 | **0.066** |
| **NIW1$_{SP}$06** | **0.040***** | **0.019** | **0.035**** | **0.021** | **0.018** | **0.024** | 0.005 | - | **0.066`** |
| **NIW2$_{SP}$06** | **0.037***** | **0.019**** | **0.026**** | **0.035**** | **0.017** | **0.020** | 0.002 | **0.030**** | - |

<0.05 **BOLD**; <0.01 **; <0.001 ***

**Table S0.2** Genetic marker discordance among *S. mentella* samples for mitochondrial control region (mtDNA clades), rhodopsin SNPs and microsatellite genotype. 'AG' indicates presence of a heterozygote. Samples in bold highlight mismatch between mtDNA and rhodopsin.

| **Sample** | **mtDNA clades** | **Rhodopsin SNP** | **Microsatellite genotype** |
|---|---|---|---|
| **IrNEdp06_29** | **A** | **A** | Deep |
| **IrWdp_36** | **B** | **G** | - |
| **IrWdp_35** | **A** | **AG** | - |
| **IrWdp_39** | **B** | **G** | - |
| **IrW$_{Int}$_12** | **A** | **AG** | - |
| **FW1dp07_7.6** | **A** | **A** | Deep |
| **FW1dp07_7.8** | A | G | Deep |
| **FW2dp07_7.15** | **A** | **A** | Deep |
| **FW2dp07_7.21** | A | - | Deep |
| **FE1sp07_8.5** | B | - | Shallow |
| **FE1sp07_8.6** | B | A | Shallow |

**Table S0.3** Pairwise $F_{ST}$ genetic differentiation among nine *S. mentella* collections. mtDNA (below diagonal) and microsatellite (above diagonal) based on 10,000 permutations. Significance tested after the False Discovery Rate (FDR) correction

| | **$IrNE_{DP}06$** | **$IrSW_{SP}06$** | **$FW1_{DP}07$** | **$FW2_{DP}07$** | **$FE1_{SP}07$** | **$FE2_{SP}07$** | **$NS_{SP}07$** | **$NIW1_{SP}06$** | **$NIW2_{SP}06$** |
|---|---|---|---|---|---|---|---|---|---|
| **$IrNE_{DP}06$** | - | **0.030***** | -0.008 | -0.004 | **0.029**** | **0.021**** | **0.029**** | **0.040***** | **0.037***** |
| **$IrSW_{SP}06$** | **0.566***** | - | **0.021** | **0.022**** | 0.002 | 0.012 | -0.003 | **0.019** | **0.019**** |
| **$FW1_{DP}07$** | **0.203**** | **0.383***** | - | -0.009 | **0.029**** | 0.017 | **0.024** | **0.035**** | **0.026**** |
| **$FW2_{DP}07$** | **0.192***** | **0.394***** | -0.080 | - | **0.022** | 0.019 | 0.010 | **0.021** | **0.035**** |
| **$FE1_{SP}07$** | **0.494***** | 0.037 | **0.214**** | **0.230**** | - | 0.007 | -0.003 | **0.018** | **0.017** |
| **$FE2_{SP}07$** | **0.635***** | 0.017 | **0.437***** | **0.445***** | 0.073 | - | 0.011 | **0.024** | **0.020** |
| **$NS_{SP}07$** | **0.628***** | **0.116** | **0.456***** | **0.464***** | **0.133** | -0.014 | - | 0.005 | 0.002 |
| **$NIW1_{SP}06$** | **0.608***** | -0.014 | **0.371***** | **0.379***** | -0.020 | 0.005 | 0.128 | - | **0.030**** |
| **$NIW2_{SP}06$** | **0.599***** | -0.032 | **0.346***** | **0.363***** | -0.020 | 0.004 | 0.071 | -0.034 | - |

<0.05 **BOLD**; <0.01 **; <0.001 ***

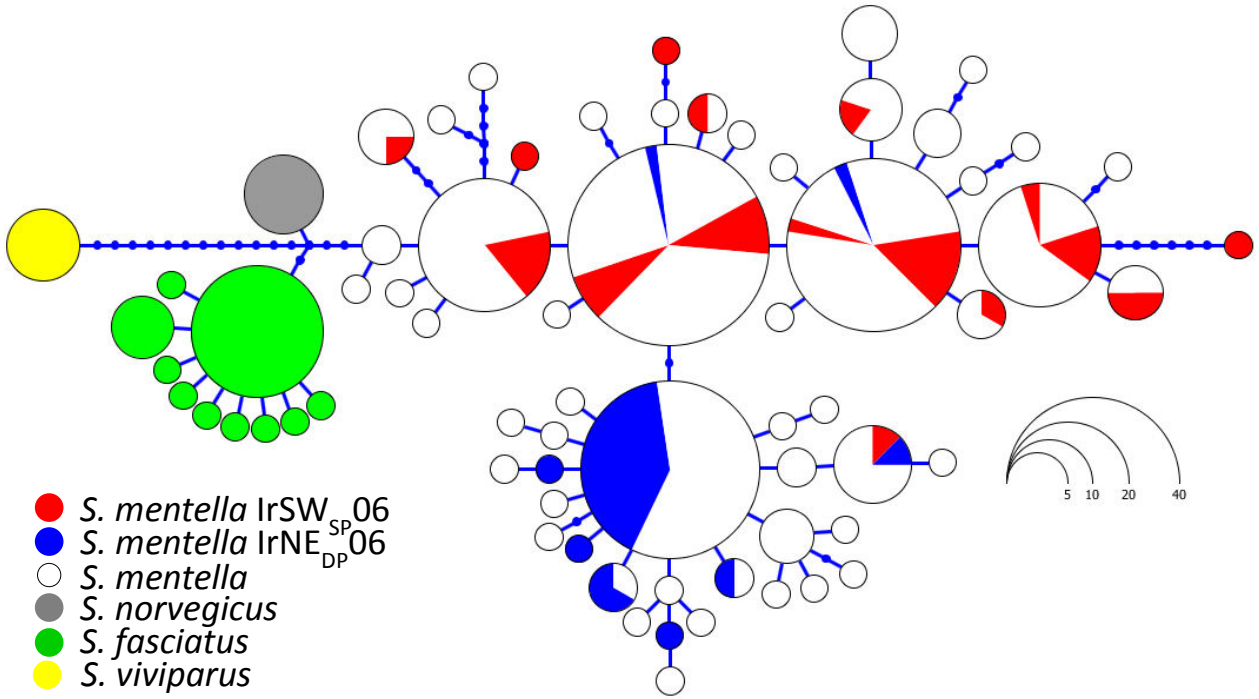



**Figure S0.1** Haplotype genealogies indicating shallow- (red) and deep-pelagic (blue) samples from Shum *et al.* (2014) within the larger mtDNA dataset generated in the present study (See Fig. 2, main article).

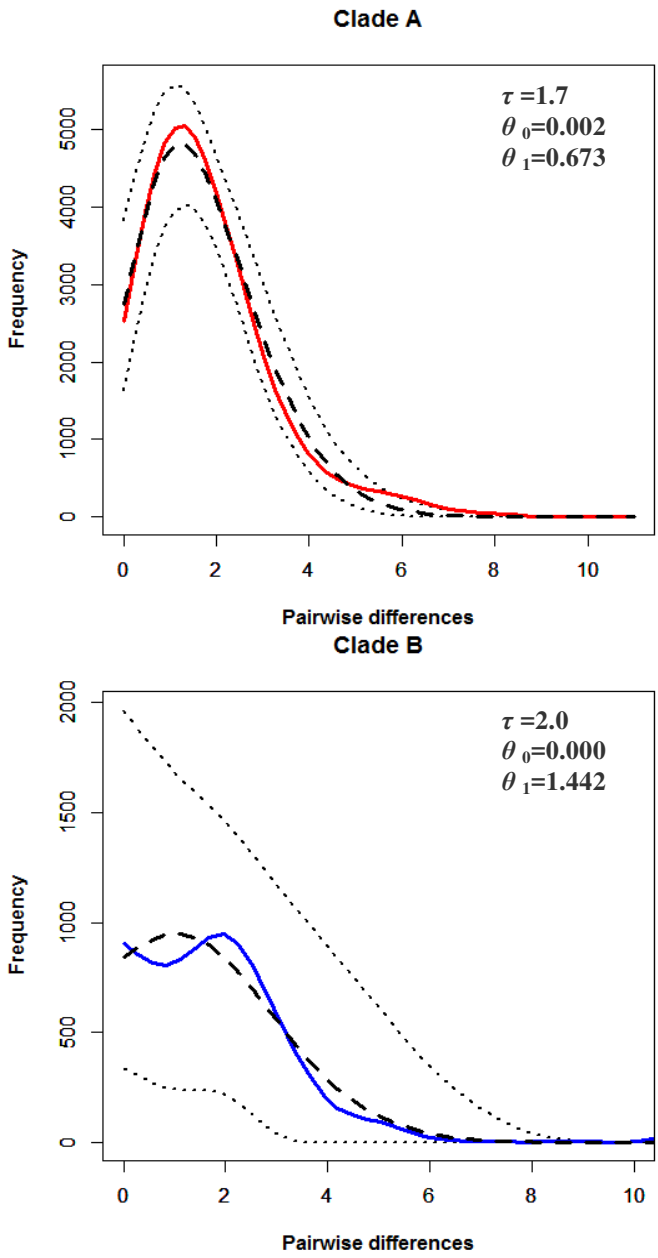



**Figure S0.2** Mismatch distributions for oceanic *S. mentella* clade A (above) and clade B (below). Red (clade A) and blue (clade B) lines represent observed pairwise difference frequency distributions; black dashed lines represent simulated distributions, and dotted curves delimit the 95% confidence limits of the simulated distributions.

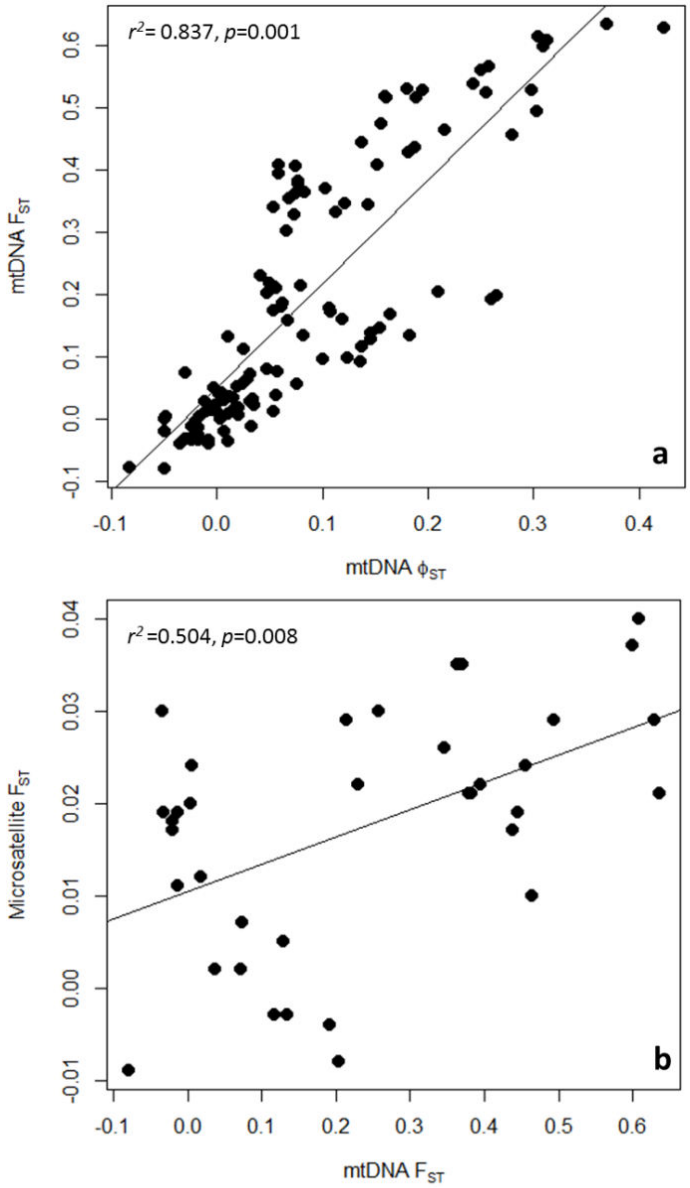



**Figure S0.3** Correlation plots between pairwise genetic distances for *S. mentella* collections from (a) 16 sampling regions based on mtDNA ($\Phi_{ST}$ vs. $F_{ST}$) and (b) 9 sampling regions based on microsatellite and mtDNA $F_{ST}$. Sampling information is provided in Table 2.1 (chapter II).

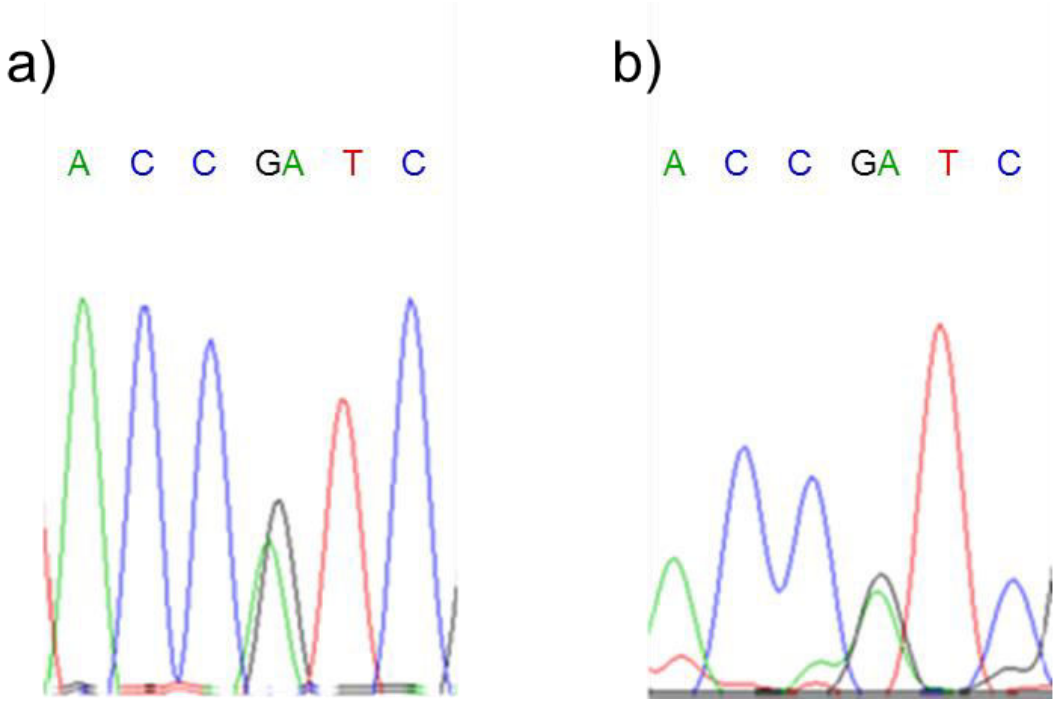



**Figure S0.4** Chromatograms showing heterozygous SNP at position 208 of the Rhodopsin gene for a) $IrW_{Int}13\_12$ and b) $IrW_{DP}13\_35$.

## References


Ackerly DD, Schwilk DW & Webb CO (2006) Niche evolution and adaptive radiation: testing the order of trait divergence. *Ecology*, **87**: S50-S61.

Afonso P, McGinty N, Graça G, Fontes J, Inácio M, Totland A & Menezes G (2014) Vertical Migrations of a Deep-Sea Fish and Its Prey. *PLoS ONE*, 9: e97884.

Alcorn SW & RJ Pascho (2002) Antibody responses by chinook salmon (*Oncorhynchus tshawytscha*) and rainbow trout (*Oncorhynchus mykiss*) to various protein antigens. *Fish & shellfish immunology*, **13**: 327-333.

Alesandrini S & Bernardi G (1999) Ancient species flocks and recent speciation events: What can rockfish teach us about cichlids (and vice versa)? *Journal of Molecular Evolution,* **49**: 814-818.

Allender CJ, Seehausen O, Knight ME, Turner GF & Maclean N (2003) Divergent selection during speciation of Lake Malawi cichlid fishes inferred from parallel radiations in nuptial coloration. *Proceedings of the National Academy of Sciences of the United States of America,* **100**: 14074-14079.

Anders S, Pyl PT & Huber W (2014) HTSeq - a Python framework to work with high-throughput sequencing data. *Bioinformatics*, **31**: 166-169.

Aris-Brosou S & Excoffier L (1996) The Impact of Population Expansion and Mutation Rate Heterogeneity on DNA Sequence Polymorphism. *Molecular Biology & Evolution,* **13**: 494-504.

Arshavsky VY (2002). Rhodopsin phosphorylation: from terminating single photon responses to photoreceptor dark adaptation. *Trends in neurosciences*, **25**: 124-126.

Arshavsky VY, Lamb TD & Pugh Jr EN (2002) G proteins and phototransduction. *Annual review of Physiology*, **64**: 153-187.

Artamonova VS, Makhrov AA, Karabanov DP, Rolskiy AY, Bakay YI & Popov VI (2013) Hybridization of beaked redfish (*Sebastes mentella*) with small redfish (*Sebastes viviparus)* and diversification of redfish (Actinopterygii: Scorpaeniformes) in the Irminger Sea. *Journal of Natural History*, **47**: 1791-1801.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM & Sherlock G (2000) Gene Ontology: tool for the unification of biology. *Nature Genetics*, **25**: 25-29.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology* **215**: 403-410.

Avise JC (2004) Molecular Markers, Natural History, and Evolution, 2nd ed. *Sunderland (Massachusetts): Sinauer Associates.*

Baeuerle PA & Henkel T (1994). Function and activation of NF-kappaB in the immune system. *Annual review of immunology*, **12**: 141-179.

Bailes HJ, Davies WL, Trezise AEO & Collin SP (2007) Visual pigments in a living fossil, the Australian lungfish *Neoceratodus forsteri*, *BMC Evolutionary Biology*, **7**:200

Baylor DA, Lamb TD & Yau KW (1979) Responses of retinal rods to single photons. *The Journal of Physiology*, **288**: 613-634.

Belkhir K, Borsa P, Chikhi L, Raufaste N & Bonhomme F (1996) GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Laboratoire Génome, Populations, Interactions, CNRS UMR 5171, Université de Montpellier II, Montpellier (France).

Benjamini Y & Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B Methodological*, **57**: 289-300.

Bergstad OA (2013) North Atlantic demersal deep-water fish distribution and biology: present knowledge and challenges for the future. *Journal of fish Biology,* **83**: 1489-1507.

Berner D, Grandchamp A-C & Hendry AP (2009) Variable progress toward ecological speciation in parapatry: stickleback across eight lake-stream transitions. *Evolution,* **63**: 1740-1753.

Beverton RJ & Holt SJ (2012) *On the dynamics of exploited fish populations* (Vol. 11). Springer Science & Business Media.

Bigg GR, Cunningham CW, Ottersen G, Pogson GH, Wadley MR & Williamson P (2008) Ice-age survival of Atlantic cod: agreement between palaeoecology models and genetics. *Proceedings of the Royal Society B,* **275**: 163-172.

Blanchon P & Shaw J (1995) Reef drowning during the last deglaciation: Evidence for catastrophic sea-level rise and ice-sheet collapse. *Geology* **23**: 4-8.

Bongaerts P, Riginos C, Hay KB, van Oppen MJ, Hoegh-Guldberg O & Dove S (2011) Adaptive divergence in a scleractinian coral: physiological adaptation of Seriatopora hystrix to shallow and deep reef habitats. *BMC evolutionary biology*, **11**: 303.

Bowen BW, Rocha LA, Toonen RJ, Karl SA & Laboratory TT (2013) The origins of tropical marine biodiversity. *Trends in Ecology and Evolution*, **28**: 359–366.

Bowmaker JK & Hunt DM (2006) Evolution of vertebrate visual pigments. *Current Biology*, **16**: 484-489.

Bowmaker JK, Semo M, Hunt DM & Jeffery G (2008) Eel visual pigments revisited: The fate of retinal cones during metamorphosis. *Visual Neuroscience*, **25**: 249–255.

Bowmaker JK, Govardovskii VI, Shukolyukov SA, Zueva LV, Hunt DM, Sideleva VG & Smirnova OG (1994) Visual pigments and the photic environment – the cottoid fish of Lake Baikal. *Vision Research* **34**: 591–605.

Brown A & Thatje S (2011) Respiratory response of the deep-sea amphipod Stephonyx biscayensis indicates bathymetric range limitation by temperature and hydrostatic pressure. *PLOS ONE* **6**: e28562.

Bruno TJ & Svoronos PD (2005) *CRC handbook of fundamental spectroscopic correlation charts*. CRC Press.

Burns ME & Baylor DA (2001) Activation, deactivation, and adaptation in vertebrate photoreceptor cells. *Annual review of neuroscience*, **24**: 779-805.

Butlin R, Debelle A, Kerth C, Snook RR, Beukeboom LW, Castillo CR, Diao W, Maan M.E, Paolucci S, Weissing FJ & van de Zande L (2012) What do we need to know about speciation?. *Trends in Ecology & Evolution*, **27**:27-39.

Cadrin SX (2000) Advances in morphometric analysis of fish stock structure. *Reviews in Fish Biology and Fisheries* **10**: 91–112.

Cadrin SX, Bernreuther M, Daníelsdóttir AK, Hjörleifsson E, Johansen T, Kerr L, Kristinsson K, Mariani S, Nedreaas K, Pampoulie C, Planque B, Reinert J, Saborido-Rey F, Sigurðsson T & Stransky C (2010) Population structure of beaked redfish, *Sebastes mentella*: evidence of divergence associated with different habitats. *ICES Journal of Marine Science,* **68**: 1617-1630.

Cadrin SX, Kerr LA & Mariani, S. eds., 2013. *Stock identification methods: applications in fishery science*. Academic Press.

Cadrin SX, Mariani S, Pampoulie C, Bernreuther M, Daníelsdóttir AK, Johansen T, Kerr L, Nedreaas K, Reinert J, Sigurðsson Þ, Stransky C (2011) Counter-comment on Cadrin *et al*. (2010) "Population structure of beaked redfish, *Sebastes mentella*: evidence of divergence associated with different habitats. ICES Journal of Marine Science, 67:1617-1630. *ICES Journal of Marine Science*, **68**: 2016-2018.

Carleton KL, Parry JWL, Bowmaker JK, Hunt DM & Seehausen OLE (2005) Colour vision and speciation in Lake Victoria cichlids of the genus *Pundamilia*. *Molecular Ecology*, **14**: 4341–4353.

Carleton KL, Hofmann CM, Klisz C, Patel Z, Chircus LM, Simenauer LH, SoodooN, Albertson RC & Ser JR (2010) Genetic basis of differential opsin gene expression in cichlid fishes. *Journal of Evolutionary Biology*, **23**: 840-853.

Chen W-J, Bonillo C & Lecointre G (2003) Repeatability of clades as a criterion of reliability: a case study for molecular phylogeny of Acanthomorpha (Teleostei) with larger number of taxa. *Molecular Phylogenetics and Evolution*, **26**: 262-288.

Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, Wohlfarth B, Mitrovica JX, Hostetler SW & McCabe AM (2009) The last glacial maximum. *Science*, **325**: 710-714.

Cornuet J-M, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Marin J-M & Estoup A (2014) DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA sequence and microsatellite data. *Bioinformatics,* **30**: 1187-1189.

Cornuet J-M, Santos F & Beaumont M (2008) Inferring population history with DIYABC: a user-friendly approach to approximate Bayesian computations. *Bioinformatics,* **24**: 2713-2719.

Craig J & Priede IG (2012) 'Bioluminescence – a source of marine energy?' *The Crown Estate*. 27 pages.

Crucifix M & Berger A (2002) Simulation of ocean-ice sheet interactions during the last deglaciation. *Paleoceanography,* **17**: 1054.

Dagorn L, Bach P & Josse E (2000) Movement patterns of large bigeye tuna (*Thunnus obesus*) in the open ocean, determined using ultrasonic telemetry. *Marine Biology*, **136**: 361–371.

Dalton BE, Lu J, Leips J, Cronin TW & Carleton KL (2015) Variable light environments induce plastic spectral tuning by regional opsin coexpression in the African cichlid fish, *Metriaclima zebra*. *Molecular ecology*, **24**: 4193-4204.

Davey JW & Blaxter ML (2010) RADSeq: next-generation population genetics. *Briefings in Functional Genomics*, **9**:416-423.

Daw T & Gray T (2005) Fisheries science and sustainability in international policy: a study of failure in the European Union's Common Fisheries Policy. *Marine Policy*, **29**: 189-197.

Doebeli M & Dieckmann U (2003) Speciation along environmental gradients. *Nature,* **421**: 259-264.

Douglas R, Djamgoz M (1990) The visual system of fish. Chapman & Hall.

Douglas RH., Mullineaux CW & Partridge JC (2000) Long-wav sensitivity in deep-sea stomiid dragonfish with far-red bioluminescence: evidence for a dietary origin of the chlorophyll-derived retinal photosensitizer of *Malacosteus niger*. *Philosophical Transactions of the Royal Society of London B* **355**: 1269-1272.

Douglas RH & Partridge JC (1997) On the visual pigments of deep-sea fish. *Journal of Fish Biology* **50**: 68–85

Douglas RH, Partridge JC & Hope AJ (1995). Visual and lenticular pigments in the eyes of demersal deep-sea fishes. *Journal of Comparative Physiology A*, **177:** 111–122.

Drevetnyak K & Nedreaas KH (2009) Historical movement pattern of juvenile beaked redfish (*Sebastes mentella* Travin) in the Barents Sea as inferred from long-term research survey series. *Marine Biology Research,* **5**: 86-100.

Drummond AJ, Suchard MA, Xie D & Rambaut A (2012) Bayesian phylogenetics with BEAUti and BEAST 1.7. *Molecular Biology and Evolution,* **29**: 1969-1973.

Dunton K (1992) Arctic biogeography: The paradox of the marine benthic fauna and flora. *Trends in Ecology & Evolution*, **7**: 183–189.

Dupanloup I, Schneider S & Excoffier L (2002) A simulated annealing approach to define the genetic structure of popuations. *Molecular Ecology,* **11**: 2571-2581.

Ebrey T & Koutalos Y (2001) Vertebrate photoreceptors. *Progress in retinal and eye research*, **20**: 49-94.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research,* **32**: 1792-1797.

EC (European Commission) (2013) EU Regulation No 1379/2013 on the common organisation of the markets in fishery and aquaculture products. *Official Journal of the European Union* **L354**: 1–21.

Engås A, Skeide R & West W (1997) The "MultiSampler": a syatem for remotely opening and closing multiple codends on a sampling trawl. *Fisheries Research,* **29**: 285-298.

Excoffier L & Licher HE (2010) Arlequin suite ver3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources,* **10**: 564-567.

FAO Fisheries and Aquaculture Department (2014) The State of World Fisheries and Aquaculture 2014. Food and Agriculture Organization of the United Nations, Rome.

Federal Agency for Agriculture and Food (2015) Directory of tradenames for fishery and aquaculture (http://www.ble.de/SharedDocs/Downloads/02_Kontrolle/02_Fischerei/01_Fischwirtschaft/HandelsbezeichnungDLat.html%3bjsessionid=FBBBED735D8259BE5D061B8C6401F9A2.1_cid335?nn=2307644)

Felsenstein J 1985 Confidence limits on phylogenies: an approach using the bootstrap. Evolution, **39**: 783–791.

Felsenstein J (1989) PHYLIP - Phylogeny Inference Package (Version 3.2). *Cladistics,* **5**: 164-166.

Filonzi L, Chiesa S, Vaghi M & Marzano FN (2010) Molecular barcoding reveals mislabelling of commercial fish products in Italy. *Food Research International*, **43**: 1383-1388.

Finn RD, Clements J & Eddy SR (2011) HMMER Web Server: Interactive Sequence Similarity Searching. *Nucleic Acids Research*, **39**: W29-37.

Finnegan AK, Griffiths AM, King RA, Machado-Schiaffino G, Porcher J-P, Garcia-Vazquez E, Bright D & Stevens JR (2013) Use of multiple markers demonstrate a cryptic western refuge and postglacial colonisation routes of Atlantic salmon (*Salmo salar* L.) in northwest Europe. *Heredity,* **111**: 1365-2540.

Flamarique NI 2013 Opsin switch reveals function of the ultraviolet cone in fish foraging. *Proceedings of the Royal Society of London B: Biological Sciences*, 280: 2012-2490.

Frank TM & Widder EA (1997) The correlation of downwelling irradiance and staggered vertical igration patterns of zooplankton in Wilkinson Basin, Gulf of Maine. *Journal of Plankton Research*, **19**: 1975-1991.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics,* **147**: 915-925.

Gaither MR, Toonen RJ, Robertson RR, Planes S &Bowen BW (2010) Genetic evaluation of marine biogeographical barriers: perspectives from two widespread Indo-Pacific snappers (*Lutjanus kasmira* and *Lutanus fulvus*). *Journal of Biogeography,* **37**: 133-147.

Galindo HM, Pfeiffer-Herbert AS, McManus MA, Chao Y, Chai F & Palumbi SR (2010) Seascape genetics along a steep cline: using genetic patterns to test predictions of marine larval dispersal. *Molecular Ecology,* **19**: 3692-3707.

Garrison T (2015) *Oceanography: an invitation to marine science*. Cengage Learning.

Gauldie RW (1991) Taking stock of genetic concepts in fisheries management. *Canadian Journal of Fisheries and Aquatic Sciences*, **48**: 722-731.

Gauthier S & Rose GA (2002) Acoustic observation of diel vertical migration and shoaling behaviour in Atlantic redfishes. *Journal of Fish Biology*, **61**: 1135–1153

Gesser B, Lund M, Lohse N, Vestergaad C, Matsushima K, Sindet-Pedersen S, Jensen SL, Thestrup-Pedersen K & Larsen CG (1996) IL-8 induces T cell chemotaxis, suppresses IL-4, and up-regulates IL-8 production by CD4+ T cells. *Journal of leukocyte biology* **59**: 407-411.

Ghosh S, May MJ & Kopp EB (1998) NF-κB and Rel proteins: evolutionarily conserved mediators of immune responses. *Annual review of immunology*, **16**: 225-260.

Govardovskii VI, Fyhrquist N, Reuter T, Kuzmin DG & Donner K (2000) In search of the visual pigment template. Visual Neuroscience, **17**: 509–528

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis Z, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren WN, Lindblad-Toh K, Friedman N &Regev A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, **29**: 644-652.

Greenwood PH (1991) Speciation. *Cichlid Fishes: behaviour, ecology and evolution*, pp.86-102.

Griffiths AM, Miller DD, Egan A, Fox J, Greenfield A & Mariani S (2013) DNA barcoding unveils skate (Chondrichthyes: Rajidae) species diversity in 'ray'-products sold across Ireland and the UK. *PeerJ*, 1:e129.

Grousset FE, Labeyrie L, Sinko JA, Cremer M, Bond G, Duprat J, Cortijo E &Huon S (1993) Patterns of ice-Rafted Detritus in the Glacial North Atlantic (40–55°N). *Paleoceanography,* **8**: 175-192.

Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. (2013) *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols*, **8**: 1494–512.

Hare MP, Guenther C & Fagan W (2005) Nonrandom larval dispersal can steepen marine clines. *Evolution,* **59**: 2509-2517.

Hardy O, Charbonnel N, Freville H & Heuertz M (2003) Microsatellites allele sizes: a simple test to assess their significance on genetic differentiation. *Genetics,* **163**: 1467-1482.

Hargrave PA (2001) Rhodopsin structure, function, and topography the Friedenwald lecture. *Investigative ophthalmology & visual science*, **42**: 3-9.

Hart N, Partridge JC, Bennett ATD & Cuthill IC (2000) Visual pigments, cone oil droplets and ocular media in four species of estridid finch. *Journal of Comparative Physiology A,* **186**: 681-694.

Hasegawa M, Kishino K & Yano T (1985) Dating the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution,* **22**: 160-174.

Hawryshyn CW, Arnold MG, Chaisson DJ & Martin PC (1989) The ontogeny of ultraviolet photosensitivity in rainbow trout (*Salmo gairdneri*). *Vision Neuroscience*, **2**: 247–254.

Helvik JV, Drivenes M, Harboe T & Seo HC (2001) Topography of different photoreceptor cell types in the larval retina of Atlantic halibut (*Hippoglossus hippoglossus*). *Journal of Experimental Biology*, **204**: 2553-2559.

Hebert PDN, Ratnasingham S & deWaard, JR (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(Suppl 1), pp.S96-S99.

Hunt DM, Dulai KS, Cowing JA, Julliot C, Mollon JD, Bowmaker JK, Li W-H. & Hewett-Emmett D (1998) Molecular evolution of trichromacy in primates. *Vision Research*, **38:** 3299–3306.

Hunt DM, Dulai KS, Partridge JC, Cottrill P & Bowmaker JK (2001) The molecular basis for spectral tuning of rod visual pigments in deep-sea fish. *Journal Experimental Biology*, **204**: 3333-3344.

Hauser L & Carvalho GR (2008) Paradigm shifts in marine fisheries genetics: ugly hypothesis slain by beautiful facts. *Fish and Fisheries,* **9**: 333-362.

Hemmer-Hansen J, Nielsen EE, Gronkjaer P & Loeschcke V (2007) Evolutionary mechanisms shaping the genetic population structure of marine fishes; lessons from the European flounder (*Platichthys flesus* L.). *Molecular Ecology,* **16**: 3104-3118.

Hemmer-Hansen J, Therkildsen NO & Pujolar, JM (2014) Population genomics of marine fishes: Next-generation prospects and challenges. *The Biological Bulletin*, **227**: 117-132.

Hedrick PW (2006) Genetic polymorphism in heterogeneous environments: the age of genomics. *Annual Review of Ecology, Evolution, and Systematics*, 67-93.

Hofmann CM, O'Quin KE, Smith A & Carleton KL (2010) Plasticity of opsin gene expression in cichlids from Lake Malawi. *Molecular Ecology*, **19**: 2064–2074.

Hutchinson WF, Culling M, Orton DC, Hänfling B, Handley LL, Hamilton-Dyer S, O'Connell TC, Richards MP & Barrett JH (2015) The globalization of naval

provisioning: ancient DNA and stable isotope analyses of stored cod from the wreck of the Mary Rose, AD 1545. *Royal Society open science*, **2**: 150199.

Hughes TP, Baird AH, Bellwood DR, Card M, Connolly SR, Folke C, Grosberg R, Hoegh-Guldberg O, Jackson JBC, Kleypas J & Lough JM (2003) Climate change, human impacts, and the resilience of coral reefs. *Science*, **301**: 929-933.

Hyde J, Kimbrell C, Budrick J, Lynn E & Vetter R (2008) Cryptic speciation in the vermilion rockfish (*Sebastes miniatus*) and the role of bathymetry in the speciation process. *Molecular Ecology,* **17**: 1122-1136.

Hyde J & Vetter RD (2007) The origin, evolution, and diversification of rockfishes of the genus *Sebastes* (Curvier). *Moleculat Phylogenetics,* **44**: 790-811.

Ingram T (2010) Speciation along a depth gradient in a marine adaptive radiation. *Proceedings of the Royal Society B,* **278**: 613-618.

Irwin DE (2012) Local adaptation along smooth ecological gradients causes phylgeographic breaks and phenotypic clustering. *The American Naturalist,* **180**: 35-49.

Jennings R, Etter R & Ficarra L (2013) Population Differentiation and Species Formation in the Deep Sea: The Potential Role of Environmental Gradients and Depth. *PLoS ONE,* **8**: e77594.

Johns GC & Avise JC (1998) A comparative summary of genetic distances in the vertebrates from the mitochondrial cytochrome b gene. *Molecular Biology and Evolution*, **15**: 1481–1490.

Jordan R, Kellogg K, Juanes F & Stauffer J (2003) Evaluation of female mate choice cues in a group of Lake Malawi mbuna (Cichlidae). *Copeia*, **2003**:181-186.

Kai Y, Nakayama K & Nakabo T (2002) Genetic differences among three colour morphotypes of the black rockfish, *Sebastes inermis*, inferred from mtDNA and AFLP analyses. *Molecular Ecology*, **11**: 2591-2598.

Kanehisa M, Goto S, Sato Y, Furumichi M & Tanabe M (2012) KEGG for integration and interpretation of largescale molecular data sets. Nucleic Acids Research, **40**: D109-D114.

Karlsson S & Mork J (2003) Selection-induced variation at the pantophysin locus (Pan I) in a Norwegian fjord population of cod (*Gadus morhua* L.). *Molecular Ecology*, **12**: 3265–3274.

Kashiyama K, Seki T, Numata H & Goto SG (2009) Molecular characterization of visual pigments in Branchiopoda and the evolution of opsins in Arthropoda. *Molecular biology and evolution*, **26**: 299-311.

Kefalov VJ (2012) Rod and Cone Visual Pigments and Phototransduction through Pharmacological, Genetic, and Physiological Approaches. *The Journal of Biological Chemistry*, **287**: 1635-1641.

Kelly RP, Oliver TA, Sivasundar A & Palumbi SR (2010) A Method for Detecting Popuation Genetic Structure in Diverse, High Gene-Flow species. *Journal of Heredity,* **101**: 423-436.

Kirchberg K, Kim TY, Möller M, Skegro D, Raju GD, Granzin J, Büldt G, Schlesinger R & Alexiev U (2011) Conformational dynamics of helix 8 in the GPCR rhodopsin controls arrestin activation in the desensitization process. *Proceedings of the National Academy of Sciences*, **108**: 18690-18695.

Knowles A & Dartnall HJA, (1977) The characterisation of visual pigments by absorption spectroscopy. In: Davson H (ed.) The Eye. Academic Press, London, pp 53-100.

Knutsen H, Jorde PE, Sannæs H, Hoelzel AR, Bergstad OA, Stefanni S, Johansen T & Stenseth NC (2009) Bathymetric barriers promoting genetic structure in the deepwater demersal fish tusk (*Brosme brosme*). *Molecular Ecology,* **18**: 3151-3162.

Kocher TD (2004) Adaptive evolution and explosive speciation: the cichlid fish model. *Nature Reviews Genetics*, **5**: 288-298.

Krogh A, Larsson B, von Heijne G & Sonnhammer ELL (2001) Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *Journal of Molecular Biology*, **305**: 567-580.

Kunz YW, Wildenburg G, Goodrich L & Callaghan E (1994) The fate of ultraviolet receptors in the retina of the atlantic salmon (*Salmo salar*). *Vision Research*, **34**: 1375-1383.

Land MF & Nilsson DE (2012) Animal eyes. Oxford University Press.

Langmead B & Salzberg S (2012) Fast gapped-read alignment with Bowtie 2. Nature Methods 9: 357-359.

Larson RJ (1980) Competition, habitat selection and the bathymetric segregation of two species of rockfish (*Sebastes*). *Ecological Monographs*, **50**: 221–239.

Lenoir J & Svenning JC (2015) Climate-related range shifts–a global multidimensional synthesis and new research directions. *Ecography*, **38**: 15-28.

Librado P & Rozas J (2009) DnaSP: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics,* **25**: 1451-1452.

Logan CA, Alter SE, Haupt AJ, Tomalty K, Palumbi SR (2008) impediment to consumer choice: overfished species are sold as Pacific red snapper. *Biological Conservation,* **141**: 1591–99.

Longmore C, Fogarty K, Neat F, Brophy D, Trueman C, Milton A & Mariani S (2010) A comparison of otolith microchemistry and otolith shape analysis for the study of spatial variation in a deep-sea teleost, *Coryphaenoides rupestris*. *Environmental biology of fishes,* **89**: 591-605.

Love MS, Yaklovich M & Thorsteinson L (2002) The Rockfish of the Northeast Pacific. *UC Press, Berkeley.*

Ludt WB & Rocha LA (2015) Shifting seas: the impacts of Pleistocene sea-level fluctuations on the evolution of tropical marine taxa. *Journal of Biogeography,* **42**: 25-38.

Lythgoe JN (1988) Light and vision in the aquatic environment. In Sensory biology of aquatic animals, Springer New York.

Lythgoe J & Partridge J (1989) Visual pigments and the acquisition of visual information. *Journal of Experimental Biology* **46**: 1-20

Magnússon J & Magnússon JV (1995) Oceanic redfish (*Sebastes mentella*) in the Irminger Sea and adjacent waters. *Scientia Marina,* **59**: 241-254.

Makhrov AA, Artamonova VS, Popov VI, Rolskiy AY, Bakay YI (2011) Comment on: Cadrin *et al*. (2010) "Population structure of beaked redfish, *Sebastes mentella*: evidence of divergence associated with different habitats. ICES Journal of Marine Science, 67: 1617–1630". *ICES Journal of Marine Science,* **68**: 2013-2015.

Malinsky M, Challis RJ, Tyers AM, Schiffels S, Terai Y, Ngatunga BP, Miska EA, Durbin R, Genner MJ & Turner GF (2015) Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. *Science*, **350**: 1493-1498.

Manfrin C, Tom M, De Moro G, Gerdol M, Giulianini PG & Pallavicini A (2015) The eyestalk transcriptome of red swamp crayfish *Procambarus clarkii*. *Gene*, **557**: 28-34.

Manousaki T, Hull PM, Kusche H, Machado-Schiaffino G, Franchini P, Harrod C, Elmer KR & Meyer A (2013) Parsing parallel evolution: ecological divergence and differential gene expression in the adaptive radiations of thick-lipped Midas cichlid fishes from Nicaragua. *Molecular ecology*, **22**: 650-669.

Marcogliese DJ, Albert E, Gagnon P & Sévigny JM (2003). Use of parasites in stock identification of the deepwater redfish (*Sebastes mentella*) in the Northwest Atlantic. *Fishery Bulletin-National Oceanic And Atmospheric Administration*, **101**: 183-188.

Mariani S & Bekkevold D (2013) The nuclear genome: neutral and adaptive markers in fisheries science. Stock Identification Methods: Applications in Fisheries Science. Elsevier, In Cadrin SX, Kerr LA & Mariani S (2013) Stock identification methods: applications in fishery science. Academic Press.

Marine Research Institute (2014) State of Marine Stocks in Icelandic Waters 2013/2014 - Prospects for the Quota Year 2014/2015. Marine Research in Iceland 176. 188 pp. [http://www.hafro.is/Bokasafn/Timarit/fjolrit-176.pdf]

Marko PB, Lee SC, Rice AM, Gramling JM, Fitzhenry T M, McAlister JS, Harper GR & Moran AL (2004) Mislabelling of a depleted reef fish. *Nature*, **430**: 309–310.

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal, **17**: 10-12.

McQuinn IH (1997) Metapopulations and the Atlantic herring. *Reviews in Fish biology and fisheries*, **7**: 297-329.

Merilä J & Crnokrak P (2001) Comparison of genetic differentiation at marker loci and quantitative traits. *Journal of Evolutionary Biology*, **14**: 892-903.

Meuthen D, Rick IP, Thünken T & Baldauf SA (2012) Visual prey detection by near-infrared cues in a fish. *Naturwissenschaften*, **99:** 1063-1066.

Michiels NK, Anthes N, Hart NS, Herler J, Meixner AJ, Schleifenbaum F, Schulte G, Siebeck UE, Sprenger D & Wucherer MF (2008) Red fluorescence in reef fish: a novel signalling mechanism? *BMC ecology*, **8**: 16.

Miller SA, Dykes DD & Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research,* **16**: 1215.

Miller D, Clarke M & Mariani S 2012 Mismatch between fish landings and market trends: A western European case study. *Fisheries Research*, 121–122: 104-114

Miller DD & Mariani S (2010) Smoke, mirrors, and mislabeled cod: poor transparency in the European seafood industry. *Frontiers in Ecology and the Environment*, **8**: 517e521.

Ministry of Industry and Ministry of Fisheries (2013) Regulations on the quality of fish and fish products (https://lovdata.no/dokument/SF/forskrift/2013-06-28-844#KAPITTEL_10).

Narum SR (2006) Beyond Bonferroni: less conservative annalyses for conservation genetics. *Conservation Genetics,* **7**: 783-787.

Nedreaas K, Johansen T & Nævdal G. (1994) Genetic studies of redfish (*Sebastes* spp.) from Icelandic and Greenland waters. *ICES Journal of Marine Science*, **51**: 461-467.

Nei M (1987) Molecular Evolutionary Genetics. *Columbia University Press, New York.*

Nei M & Tajima F (1981) DNA polymorphism detectable by restriction endonucleases. *Genetics,* **97**: 145-163.

Nicolè S, Negrisolo E, Eccher G, Mantovani R, Patarnello T, Erickson DL, Kress WJ & Barcaccia G (2012) DNA barcoding as a reliable method for the authentication of commercial seafood products. *Food Technology and Biotechnology*, **50**: 387-398.

Nielsen EE, Cariani A, Mac Aoidh E, Maes GE, Milano I, Ogden R, Taylor M, Hemmer-Hansen J, Babbucci M, Bargelloni L & Bekkevold D (2012) Gene-associated markers provide tools for tackling illegal fishing and false eco-certification. *Nature Communications*, **3**: 851.

Nielsen EE, Hansen MM & Meldrup D (2006) Evidence of microsatellite hitch-hiking selection in Atlantic cod (*Gadus morhua* L.): implications for inferring population structure in nonmodel organisms. *Molecular Ecology*, **15**: 3219-3229.

Nielsen EE, Hemmer-Hansen J, Larsen PF & Bekkevold D (2009) Population genomics of marine fishes: identifying adaptive variation in space and time. *Molecular ecology*, **18**: 3128-3150.

Nosil P (2012) Ecological Speciation, 1st edn. *Oxfor University Press, New York.*

Nosil P, Funk DJ & Ortiz-Barrientos D (2009a) Divergent selection and heterogeneous genomic divergence. *Molecular ecology*, **18**: 375-402.

Nosil P, Harmon LJ & Seehausen O (2009b) Ecological explanations for (incomplete) speciation. *Trends in Ecology and Evolution*, **24**: 145–156.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH & Wagner H (2011) vegan: Community Ecology Package. R package version 2.0-2.

Olsen SM, Hansen B, Quadfasel D & Osterhus S (2008) Observed and modelled stability of overflow across the Greenland-Scotland ridge. *Nature,* **455**: 519-522.

Ou WB, Yi T, Kim JM & Khorana HG (2011) The roles of transmembrane domain Helix-III during rhodopsin photoactivation. *PLoS ONE* **6**:e17398.

Palumbi SR (1994) Genetic divergence, reproductive isolation, and marine speciation. Annual Review of Ecology and Systematics, 547-572.

Pampoulie C & Daníelsdóttir AK (2008) Resolving species identification problems in the genus *Sebastes* using nuclear genetic markers. *Fisheries Research,* **93**: 54-63

Pampoulie C, Ruzzante DE, Chosson V, Jörundsdóttir TD, Taylor L, Thorsteinsson V, Daníelsdóttir AK & Marteinsdóttir G (2006) The genetic structure of Atlantic cod (*Gadus morhua*) around Iceland: insight from microsatellites, the Pan I locus, and tagging experiments. *Canadian Journal of Fisheries and Aquatic Sciences*, **63**: 2660-2674.

Pampoulie C, Slotte A, Óskarsson GJ, Helyar SJ, Jónsson A, Ólafsdóttir G, Skírnisdóttir S, Libungan LA, Jacobsen JA, Joensen H, Nielsen HH, Sigurðsson SK, Daníelsdóttir AK (2015) Stock structure of Atlantic herring *Clupea harengus* in the Norwegian Sea and adjacent waters. *Marine Ecology Progress Series*, **522**: 219-230.

Partridge JC, Archer SN. & Lythgoe JN (1988) Visual pigments in the individual rods of deep-sea fish. *Journal of Comparative Physiology A*, **162**: 543-550.

Patarnello T, Volckaert F & Castilho R (2007) Pillars of Hercules: Is the Atlantic-Mediterranean transition a phylogeographical break? *Molecular Ecology,* **16**: 4426-4444.

Pauly D & Zeller D (Editors) (2015) Sea Around Us Concepts, Design and Data (seaaroundus.org)

Pavey SA, Collin H, Nosil P & Rogers SM (2010) The role of gene expression in ecological speciation. *Annals of the New York Academy of Sciences*, **1206**: 110-129.

Peakall R & Smouse PE (2006) Genalex 6: genetic analysis in Excel. population genetic software for teaching and research. *Molecular Ecology Notes,* **6**: 288-295.

Pedchenko AP (2005) The role of interannual environmental variations in the geographic range of spawning and feeding concentrations of redfish *Sebastes mentella* in the Irminger Sea. *ICES Journal of Marine Science,* **62**: 1501-1510.

Perry AL, Low PJ, Ellis JR & Reynolds JD (2005) Climate change and distribution shifts in marine fishes. *Science*, **308:** 1912-1915.

Petersen TN, Brunak S, von Heijne G & Nielsen H (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, **8**: 785-786.

Petit R & Excoffier L (2009) Gene flow and species delimitation. *Trends in Ecology and Evolution,* **24**: 386-393.

Pétursdóttir H, Gislason A, Falk-Petersen S, Hop H & Svavarsson J (2008) Trophic interactions of the pelagic ecosystem over the Reykjanes Ridge as evaluated by fatty acid and stable isotope analyses. *Deep Sea Research II* **55**:83–93.

Phillips GA, Carleton KL & Marshall NJ (2015) Multiple genetic mechanisms contribute to visual sensitivity variation in the Labridae. *Molecular biology and evolution*, msv213.

Pikanowski RA, Morse WW, Berrien PL, Johnson DL, McMillan DG (1999) Redfish, *Sebastes* spp., life history and habitat characteristics. *NOAA Technical Memorandum NMFS-NE-132*, 19pp

Planque B, Kristinsson K, Astakhov A, Bernreuther M, Bethke E, Drevetnyak K, Nedreaas K, Reinert J, Rolskiy A & Sigurðsson T (2013) Monitoring beaked redfish (*Sebastes mentell*a) in the North Atlantic, current challenges and future prospects. *Aquatic Living Resources,* **26**: 293-306.

Posada D & Crandall KA (1998) MODELTEST: testing the model of DNA substitution. *Bioinformatics,* **14**: 817-818.

Powell S, Szklarczyk D, Trachana K, Roth A, Kuhn M, Muller J, Arnold R, Rattei T, Letunic I, Doerks T, Jensen LJ, von Mering C & Bork P (2012) eggNOG v3.0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges. *Nucleic Acids Research*, **40**: D284-D289.

Power DJ & Ni IH (1985) Morphometric differences between golden redfish (*Sebastes* marinus) and beaked redfish (*S. mentella* and S. fasciatus). *J. North Atlant. Fish. Sci.* **6**:1-7.

Premont RT, Inglese J & Lefkowitz RJ (1995). Protein kinases that phosphorylate activated G protein-coupled receptors. *The FASEB Journal*, **9**: 175-182.

Punta PC, Coggill RY, Eberhardt J, Mistry J, Tate C, Boursnell N, Pang K, Ceric FJ, Clements A, Heger L, Holm ELL, Sonnhammer SR, Eddy A & Bateman RD (2012) The Pfam protein families database. *Nucleic Acids Research*, **40**: D290-D301.

Purves D, Augustine GJ, Fitzpatrick D, Katz LC, LaMantia A-S, McNamara JO & Williams SM (2001) Neuroscience, 5th edition Sunderland (MA): Sinauer Associates.

Rambaut A, Suchard M, Xie D & Drummond A (2014) Traver v1.6, Available from http://beast.bio.ed.ac.uk/Tracer.

Ranz JM & Machado CA (2006) Uncovering evolutionary patterns of gene expression using microarrays. *Trends in Ecology & Evolution*, **21**: 29-37.

Ratnasingham S & Hebert PDN (2007) BOLD: the barcode of life data system. Molecular Ecology Notes **7**: 355–64.

Raymo ME, Ruddiman WF, Shackleton NJ & Oppo DW (1990) Evolution of Atlantic-Pacific $\delta^{13}C$ gradients over the last 2.5 m.y. *Earth and Planetary Science Letters,* **97**: 353-368.

Rehbein H (2013) Differentiation of fish species by PCR-based DNA analysis of nuclear genes. *European Food Research and Technology*, **236**: 979-990.

Renninger SL, Gesemann M & Neuhauss SC (2011) Cone arrestin confers cone vision of high temporal resolution in zebrafish larvae. *European Journal of Neuroscience*, **33**: 658–667.

Riginos C, Buckley Y, Blomberg S & Treml E (2014) Dispersal capacity predicts both population genetic structure and species richness in reef fishes. *The american Naturalist,* **184**: 52-64.

Ringelberg J (1995) Changes in light intensity and diel vertical migration: a comparison of marine and freshwater environments. *Journal of the Marine Biological Association of the United Kingdom*, **75**: 15-25.

Robinson MD, McCarthy DJ & Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**: 139-40.

Rocha-Olivares A &Vetter RD (1999) Effects of oceanographic circulation on the gene flow, genetic structure, and phylogeography of the rosethorn rockfish (*Sebastes helvomaculatus*). *Canadian Journal of Fisheries and Aquatic Sciences*, **56**: 803–813.

Roelofs D, Janssens TK, Timmermans MJ, Nota B, Mariën J, Bochdanovits Z, Ylstra B & Van Straalen NM (2009) Adaptive differences in gene expression associated with heavy metal tolerance in the soil arthropod *Orchesella cincta*. *Molecular Ecology*, **18**: 3227-3239.

Rogers AD (2000) The role of the oceanic oxygen minima in generating biodiversity in the deep sea. *Deep-Sea Research Part II: Topical Studies in Oceanography*, **47**: 119–148.

Roques S, Sévigny J-M & Bernatchez L (2002) Genetic structure of deep-water redfish, *Sebastes mentella*, populations across the North Atlantic. *Marine Biology*, **140**: 297-307.

Roy D, Hurlbut TR & Ruzzante DE (2012) Biocomplexity in a demersal exploited fish, white hake (*Urophycis tenuis*): depth-related structure and inadequacy of current management approaches. *Canadian Journal of Fisheries and Aquatic Sciences,* **69**: 415-429.

Rubec PJ, McGlade JM, Trottier BL & Ferron A (1991) Evaluation for methods for separation of Gulf of Saint-Lawrence beaked redfishes *Sebastes* fasciatus and *S.*

*mentella*: malate dehydrogenase mobility patterns compared with extrinsic gasbladder muscle passages and anal fin ray counts. *Canadian Journal of Fisheries and Aquatic Sciences*, **48**: 640–660.

Ruzzante DE, Mariani S, Bekkevold D, André C, Mosegaard H, Clausen LA, Dahlgren TG, Hutchinson WF, Hatfield EM, Torstensen E & Brigham J (2006) Biocomplexity in a highly migratory pelagic marine fish, Atlantic herring. *Proceedings of the Royal Society of London B: Biological Sciences*, **273**: 1459-1464.

Sabbah S, Laria R, Gray S & Hawryshyn C (2010) Functional diversity in the color vision of cichlid fishes. *BMC Biology*, **8**: 133.

Saborido-Rey F, Garabana D & Cerviño S (2004) Age and growth of redfish (*Sebastes marinus*, *S. mentella*, and *S. fasciatus*) on the Flemish Cap (Northwest Atlantic). *ICES Journal of Marine Science*, **61**: 231-242.

Sala-Bozano M, Ketmaier V, Mariani S (2009) Contrasting signals from multiple markers illuminate population connectivity in a marine fish. *Molecular Ecology,* **18**: 4811–4826.

Sallese M, Mariggiò S, Collodel G, Moretti E, Piomboni P, Baccetti B & De Blasi A (1997) G Protein-coupled Receptor Kinase grk4 molecular analysis of the four isoforms and ultrastructural localization in spermatozoa and germinal cells. *Journal of Biological Chemistry*, **272**: 10188-10195.

Salzburger W, Ewing GB & Haeseler A (2011) The performance of phylogenetic algorithms in estimating haplotype genealogies with migration. *Molecular Ecology,* **20**: 1952-1963.

Saitou N & Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**: 406-425.

Seehausen O, Terai Y, Magalhaes IS, Carleton KL, Mrosso HD, Miyagi R, van der Sluijs I, Schneider MV, Maan ME, Tachida H & Imai H (2008) Speciation through sensory drive in cichlid fish. *Nature*, ***455***: 620-626.

Secor DH, (2013) The unit stock concept: bounded fish and fisheries. Stock Identification Methods: Applications in Fisheries Science. Elsevier, In Cadrin SX, Kerr LA & Mariani S eds., (2013) Stock identification methods: applications in fishery science. Academic Press.

Selkoe KA & Toonen RJ (2011) Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology Progress Series,* **436**: 291-305.

Serra-Pereira B, Moura T, Griffiths AM, Gordo LS, Figueiredo I (2010) Molecular barcoding of skates (Chondrichthyes: Rajidae) from the southern Northeast Atlantic. *Zoologica Scripta*, **40**: n76-84.

Shum P, Pampoulie C, Kristinsson K & Mariani S (2015) Three-Dimensional Post-Glacial Expansion And Diversification Of An Exploited Oceanic Fish, *Molecular Ecology*, **24**: 3652–3667.

Shum P, Pampoulie C, Sacchi C & Mariani S (2014) Divergence by depth in an oceanic fish. *PeerJ,* **2**: e525.

Silva G, Horne JB & Castilho R (2014) Anchovies go north and west without losing diversity: post-glacial range expansions in a small pelagic fish. *Journal of Biogeography,* **41**: 1171-1182.

Sivasundar A & Palumbi S (2010) Parallel amino acid replacements in the rhodopsin of the rockfishes (*Sebastes* spp.) associated with shifts in habitat depth. *Journal of Evolutionary Biology,* **23**: 1195-1169.

Smith SO (2010) Structure and activation of the visual pigment rhodopsin. *Annual Review of Biophysics* **39**: 309-328

Smith, E. J., J. C. Partridge, K. N. Parsons, E. M. White, I. C. Cuthill, A. T. D. Bennett & S. C. Church (2002) Ultraviolet vision and mate choice in the guppy (*Poecilia reticulata*). *Behavioral Ecology* **13**:11–19.

Smolka J, Zeil J & Hemmi JM (2011) Natural visual cues eliciting predator avoidance in fiddler crabs. *Proceedings of the Royal Society of London B*, **278**: 3584-3592.

Solem ST & Stenvik J (2006) Antibody repertoire development in teleosts - a review with emphasis on salmonids and *Gadus morhua* L. *Developmental & Comparative Immunology*, **30**: 57-76.

Somero GN (1992) Adaptations to high hydrostatic pressure. *Annual Review of Physiology,* **54**: 557-577.

Sommer S (2005) The importance of immune gene variability (MHC) in evolutionary ecology and conservation. *Frontiers in Zoology*, **2**: 16.

Sousa-Santos C, Robalo JI, Collares-Pereira M-J & Almada VC (2005) Heterozygous indels as useful tools in the reconstruction of DNA sequences and in the assessment of ploidy level and genomic constitution of hybrid organisms. *DNA Sequence*, **16**: 462-467.

Spady TC, Seehausen O, Loew ER, Jordan RC, Kocher TD & Carleton KL (2005) Adaptive molecular evolution in the opsin genes of rapidly speciating cichlid species. *Molecular Biology and Evolution*, **22**: 1412–1422.

Stefánsson MÖ, Reinert J, Sigurđsson T, Kristinsson K, Nedreaas K & Pampoulie C (2009b) Depth as a potential driver of genetic structure of *Sebastes mentella* across the North Atlantic Ocean. *ICES Journal of Marine Science,* **66**: 680-690.

Stefánsson MÖ, Sigurđsson T, Pampoulie C, Daníelsdóttir AK, Thorgilsson B, Ragnarsdóttir R, Gíslason D, Coughlan J, Cross TF, Bernatchez L (2009a) Pleistocene genetic legacy suggests incipient species of *Sebastes mentella* in the Irminger Sea. *Heredity,* **102**: 514-524.

Steinke D, Zemlak TS, Boutillier JA & Hebert PDN. (2009) DNA barcoding Pacific Canada's fishes. *Marine Biology*, 156: 2641–2647.

Stevens JD, Bradford RW & West GJ (2010) Satellite tagging of blue sharks (*Prionace glauca*) and other pelagic sharks off eastern Australia: depth behaviour, temperature experience and movements. *Marine biology*, **157**: 575–591.

Strand TM, Segelbacher G, Quintela M, Xiao L, Axelsson T & Höglund J (2012) Can balancing selection on MHC loci counteract genetic drift in small fragmented populations of black grouse? *Ecology and evolution*, **2**: 341-353.

Stransky C (2005) Geographic variation of golden redfish (*Sebastes marinus*) and deep-sea redfish (*S. mentella*) in the North Atlantic based on otolith shape analysis. *ICES Journal of Marine Science,* **62**: 1691–1698.

Stransky C, Gudmundsdóttir S, Sigurðsson T, Lemvig S, Nedreaas K & Saborido-Rey F (2005a) Age determination and growth of Atlantic redfish (*Sebastes marinus* and *S. mentella*): bias and precision of age readers and otolith preparation methods. *ICES Journal of Marine Science,* **62**: 655-670.

Stransky C, Kanisch G, Kruger A & Purkl S (2005b) Radiometric age validation of golden redfish (*Sebastes marinus*) and deep-sea redfish (*S. mentella*) in the Northeast Atlantic. *Fisheries Research,* **74**: 186-197.

Tajima F (1983) Evolutionary relationship of DNA-sequences in finite populations. *Genetics,* **105**: 437-460.

Tamura K, Stecher G, Peterson D, Filipski A & Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Molecular Biology and Evolution, **30**: 2725-2729.

Team RDC (2005) R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing, Vienna.*

Takechi M & Kawamura S (2005) Temporal and spatial changes in the expression pattern of multiple red and green subtype opsin genes during zebrafish development. *Journal of Experimental Biology*, **208**: 1337–1345.

Taylor SM, Loew ER & Grace MS (2011) Developmental shifts in functional morphology of the retina in Atlantic tarpon, *Megalops atlanticus* (Elopomorpha: Teleostei) between four ecologically distinct life-history stages. *Visual Neuroscience*, **28**: 309–323.

Therkildsen NO, Hemmer-Hansen J, Hedeholm RB, Wisz MS, Pampoulie C, Meldrup D, Bonanomi S, Retzel A, Olsen SM & Nielsen EE (2013) Spatiotemporal SNP analysis reveals pronounced biocomplexity at the northern range margin of Atlantic cod *Gadus morhua*. *Evolutionary Applications,* **6**: 690-705.

Trapnell C, Pachter L & Salzberg (2009) SL: TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*, **25**: 1105-1111.

Trujillo AP & Thurman HV (2008) *Essentials of oceanography*. Upper Saddle River: Pearson Prentice Hall.

Tsuchi R (1997) Marine climatic responses to Neogene tectonics of the Pacific ocean seaways. *Tectonophysics*, **281**: 113–124

Turner JR, White EM, Collins MA, Partridge JC & Douglas RH (2009) Vision in lanternfish (Myctophidae): adaptations for viewing bioluminescence in the deep-sea. *Deep Sea Research I* **56**:1003–17.

Valen R, Edvardsen RB, Søviknes AM, Drivenes Ø & Helvik JV (2013) Molecular evidence that only two opsin subfamilies, the blue light- (SWS2) and green light-sensitive (RH2), drive color vision in Atlantic cod (*Gadus morhua*). *PLos One* **9**, e115436

Valentin AE, Penin X, Chanut JP, Power D & Sévigny JM, (2014) Combining microsatellites and geometric morphometrics for the study of redfish (*Sebastes* spp.) population structure in the Northwest Atlantic. *Fisheries Research*, **154**: 102-119.

Venables WN & Ripley BD (2002) Modern Applied Statistics with S. *Fourth Edition. Springer, New York.*

Venerus LA, Ciancio JE, Riva-Rossi C, Gilbert-Horvath EA, Gosztonyi AE & Garza JC (2013) Genetic structure and different color morphotypes suggest the occurrence and bathymetric segregation of two incipient species of *Sebastes* off Argentina. *Naturwissenschaften*, **100**: 645-658.

Viñas J & Tudela S (2009) A validated methodology for genetic identification of tuna species (genus Thunnus). *PLoS ONE,* **4**: e7606.

Vonlanthen P, Roy D, Hudson AG, Largiader CR, Bittner D & Seehausen O (2008) Divergence along a steep ecological gradient in lake whitefish (*Coregonus* sp.). *Journal of Evolutionary Biology,* Issue **22**: 498-514.

Wang Z, Gerstein M & Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, **10**: 57-63.

Waples RS (1998) Separating the Wheat From the Chaff: Patterns of Genetic Differentiation in High Gene Flow Species. *Journal of Heredity,* **89**: 438-450.

Waples RS & Gaggiotti (2006) INVITED REVIEW: What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, **15**: 1419-1439.

Ward RD, Zemlak TS, Innes BH, Last PR & Hebert PD (2005) DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **360**: 1847–1857.

Warrant E J & Locket N (2004) Vision in the deep sea. *Biological Reviews,* **79**, 671-712.

Watson RA, Green BS, Tracey SR, Farmery A & Pitcher T 2015 Provenance of global seafood. *Fish and Fisheries*, DOI: 10.1111/faf.12129.

Weersing KA & Toonen (2009) Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series*, **393**: 1-12.

Wegner KM, Reusch TBH & Kalbe M (2003) Multiple parasites are driving major histocompatibility complex polymorphism in the wild. *Journal of evolutionary biology*, **16**: 224-232.

Whitehead A & Crawford DL (2006) Variation within and among species in gene expression: raw material for evolution. *Molecular Ecology*, **15**: 1197-1211.

Widder EA, Greene CH & Youngbluth MJ (1992) Bioluminescence of sound-scattering layers in the Gulf of Maine. *Journal of Plankton Research* **14**:1607–24.

Widder EA (2002) Bioluminescence and the pelagic visual environment. *Marine and Freshwater Behaviour and Physiology*, **35**: 1–26.

Widder EA (2010) Bioluminescence in the ocean: origins of biological, chemical, and ecological diversity. *Science*, **328**: 704–708.

Wong EHK & Hanner R. 2008. DNA barcoding detects market substitution in North American seafood. *Food Research International*, **41**: 828-37.

Yancey PH, Gerringer ME, Drazen JC, Rowden AA & Jamieson AJ (2014) Marine fish may be biochemically constrained from inhabiting the deepest ocean depths. *Proceedings of the National Academy of Sciences,* **111**: 4461-4465.

Ye J, Fang L, Zheng H, Zhang Y, Chen J, Zhang Z, Wang J, Li S, Li R, Bolund L & Wang J (2006) WEGO: a web tool for plotting GO annotations. *Nucleic acids research*, **34**: W293-W297.

Yokoyama S (2000) Molecular evolution of vertebrate visual pigments. *Progress in Retinal and Eye Research* **19**: 385-419

Yokoyama S (2002) Molecular evolution of color vision in vertebrates. *Gene* **300**: 69-78.

Yokoyama S, Tada T, Zhang H & Britt L (2008) Elucidation of phenotypic adaptations: molecular analyses of dim-light vision proteins in vertebrates. *Proceedings of the National Academy of Science of the United States of America* **105**: 13480-13485

Yokoyama S & Takenaka N (2004) The molecular basis of adaptive evolution of squirrelfish rhodopsins. *Molecular biology and evolution* **21**: 2071–2078.

Zhu X, Brown B, Li A, Mears AJ, Swaroop A & Craft CM (2003) GRK1-dependent phosphorylation of S and M opsins and their binding to cone arrestin during cone phototransduction in the mouse retina. *Journal of Neuroscience*, **23**: 6152–6160.